

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 27, 2004, 21:22:32 ; Search time 5635 Seconds
(without alignments)
4684.278 Million cell updates/sec

Title: us-09-890-475-1
Perfect score: 3104
Sequence: 1 MSNVPPTVAAQPTTANPLL.....RYLGSLNQRSPRNSSLDPK 609

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 segs, 2167151695 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=x1p
-Q=/cgm2_1/USPRO_epool_p/US09890475/runac_24022004_135311_5291/app_query.fasta_1.775
-DB=genbml -QMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blousum62 -TRANS=human40.cdd -LIST=45
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTMT=ptco -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09890475 @CGN_1_1_4958 @runac_24022004_135311_5291 -NCPU=6 -ICPU=3
-NO_WMAP -LARGESTERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Genbml:*
1: gb_ba:*
2: gb_ncg:*
3: gb_in:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sbs:*
27: em_un:*
28: em_vl:*

29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pin:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3104	100.0	1830	8	AF228500	Arabiidops
2	3104	100.0	2257	6	AX032763	Sequence
3	2981	96.0	3761	6	AX032762	Sequence
4	2981	96.0	3761	8	AF228499	Arabiidops
5	2975	95.8	3150	8	AY198403	Arabiidops
6	2975	95.8	3261	8	AY198402	Arabiidops
7	2975	95.8	3739	8	AY198401	Arabiidops
8	2969	95.7	3301	8	AY198405	Arabiidops
9	2960	95.0	3758	8	AY198404	Arabiidops
10	2936	94.6	91849	8	FEW23	Arabiidops
11	2936	94.6	197975	8	ATCHRIV2	Arabiidops
12	931	30.0	558	8	AY092681	Arabiidops
13	931	30.0	558	8	AY092683	Arabiidops
14	931	30.0	558	8	AY092691	Arabiidops
15	931	30.0	558	8	AY092693	Arabiidops
16	927	29.9	558	8	AY092677	Arabiidops
17	927	29.9	558	8	AY092679	Arabiidops
18	927	29.9	558	8	AY092694	Arabiidops
19	927	29.9	558	8	AY092695	Arabiidops
20	927	29.9	558	8	AY092696	Arabiidops
21	921	29.7	558	8	AY092678	Arabiidops
22	921	29.7	558	8	AY092680	Arabiidops
23	921	29.7	558	8	AY092682	Arabiidops
24	921	29.7	558	8	AY092684	Arabiidops
25	921	29.7	558	8	AY092685	Arabiidops
26	921	29.7	558	8	AY092686	Arabiidops
27	921	29.7	558	8	AY092687	Arabiidops
28	921	29.7	558	8	AY092688	Arabiidops
29	921	29.7	558	8	AY092689	Arabiidops
30	921	29.7	558	8	AY092690	Arabiidops
31	921	29.7	558	8	AY092692	Arabiidops
32	845	27.2	511	8	AY176673	Brassica
33	658	21.2	503	8	AY092437	Arabiidops
34	658	21.2	503	8	AY092438	Arabiidops
35	658	21.2	503	8	AY092439	Arabiidops
36	658	21.2	503	8	AY092440	Arabiidops
37	658	21.2	503	8	AY092441	Arabiidops
38	658	21.2	503	8	AY092442	Arabiidops
39	658	21.2	503	8	AY092443	Arabiidops
40	658	21.2	503	8	AY092444	Arabiidops
41	658	21.2	503	8	AY092445	Arabiidops
42	658	21.2	503	8	AY092446	Arabiidops
43	658	21.2	503	8	AY092447	Arabiidops
44	658	21.2	503	8	AY092448	Arabiidops
45	658	21.2	503	8	AY092449	Arabiidops

RESULT 1

ALIGNMENTS

AF228500 1830 bp mRNA linear PLN 14-OCT-2000
 LOCUS AF228500
 DEFINITION Arabidopsis thaliana FRIGIDA (FRI) mRNA, complete cds.
 ACCESSION AF228500
 VERSION AF228500.1 GI:10801175
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 1830)
 Johnson, U., West, J., Lister, C., Michaels, S., Amadio, R. and
 Dean, C.
 TITLE Molecular analysis of FRIGIDA, a major determinant of natural
 variation in Arabidopsis flowering time
 JOURNAL Science 290 (5490), 344-347 (2000)
 MEDLINE 20485641
 PUBMED 11030654
 REFERENCE 2 (bases 1 to 1830)
 Johnson, U. and Lister, C.
 TITLE Direct Submission
 JOURNAL Submitted (25-JAN-2000) Molecular Genetics, John Innes Centre,
 Colney Lane, Norwich, Norfolk NR4 7UH, U.K.
 FEATURES
 source
 1. 1830
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /cultivar="ecotype H51"
 /db_xref="taxon:3702"
 /chromosome="4"
 /map="between m151 and m122"
 1. 1830
 /gene="FRI"
 1. 1830
 /gene="FRI"
 /note="late flowering gene"
 /codon_start=1
 /product="FRIGIDA"
 /protein_id="AAG23415.1"
 /db_xref="GI:10801175"
 /translation="MSNYPPYTAQPTTANPLIORHSEPRREELPKIVTESTSD
 ITIGSKQPOLFKSIDELAASVAVETTKRQDDIQRKESIBNADIKLESNEYLA
 ANNNFHQPMSPFRNNVSVETTVTSPPSQETVETSNKRGGMKCSKGLRKT
 IYANISDAKIMEELPSALKAKEPAKFLVDTIGFYQGRAPTKESPMASAVSL
 ILIBSFLMPDRKGVKXISWIKDSEATAVAMRKIVMGIGLAAEKMDARGILL
 VACFVPSNFRSTDLDLIRMSGSNBIAGALKRSQFLVPMYSGIVSEIKRGMIEAL
 EVVYTFGEMDKSALVLTSLFKMSKESPERAKRQSLPAFKKATKQLAVLSVNO
 CMETHKLDPAKELPGMOIKBOIVLEKDTLIDKEMERKASLIMEBALAKRMNG
 QIKRRRLSPMENPPTVSSSYSPITPRDRSPSGRDDDOEIKALVSYIGPSTPHRS
 RSPRTVPLPHFGIGRSVYAEHLAPNSYSPGHGRLLHRQYSPSLVHGQRHPLQYSP
 PHFGQQLPYGIQVRHSPSEERYLGLSNRSPRSNSLDPK"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2, 77e-195 Length: 1830
 Score: 3104.00 Matches: 609
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 8 Gaps: 0
 US-09-890-475-1 (1-609) x AF228500 (1-1830)
 QY 1 MetSerAsnTyrProProThrValAlaAlaGlnProThrThrThAlaAsnProLeuLeu 20
 Db 1 ATGTCCAATTATCCACCGAGGTGGGGGCAACCCACACGACGCAATCCACTGTG 60
 QY 21 GlnATGhIeGlnSerGlnGlnArgArgGlnLeuProTyrIleValGlnThrGlnSer 40
 Db 61 CAGCGCAATCAATCTGGAACGACGAGAGAAATTCGGAAGATTGTGAAAACAGAGCT 120
 QY 41 ThrSerMetAspIleThrIleGlnSerIysGlnProGlnPheLeuIysSerIleAsp 60

Db 121 ACAAGTAGACATTCGATCGCTCAATCTAAGACAGCCCTCAATTTTGAATCCATAGAC 180
 QY 61 GlnLeuAlaAlaPheSerValAlaValAlaGlnThrPheIysArgGlnPheAspLeuGln 80
 Db 181 GAATTAGCTGCTTTTCAGTTGAGTGAAGAAACCTTCAACGCCCATTCATCTTCAG 240
 QY 81 LysHisIleGlnSerIleGlnAsnAlaIleAspSerIysLeuGlnSerAsnGlnValVal 100
 Db 241 AACCAACATCAGTCAATCGAAACGCAATTCGAACTGAGAGTAAACGCGTTGTC 300
 QY 101 LeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerProAspArgAsnVal 120
 Db 301 CTGGCGCGCGGACAAATATTTTCATCGACCGCATTTATCGCTCCGGGAAATATGA 360
 QY 121 SerValGlnThrThyValThrValSerGlnProSerGlnGlnIleValProGlnThrSer 140
 Db 361 TCTGTAGAAACCAACGCTCACTGAGCCCAACGCTCAGAGATTGTACCGAGAGCTGC 420
 QY 141 AsnIysProGlnGlnGlnIysArgMetCysGlnLeuMetCysSerIysGlnLeuArgIysTyr 160
 Db 421 AATTAACCGAGGGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
 QY 161 IleTyrAlaAsnIleSerAspGlnAlaIysLeuMetGlnGlnIleProSerAlaLeuIys 180
 Db 481 ATATAGCGCAATATCTCTGATCAAGCTAATGTAAGAGATTCCTTCAGCTTGAAA 540
 QY 181 LeuAlaIysGlnProAlaIysPheValLeuAspCysIleGlnIysPheTyrLeuGlnIys 200
 Db 541 TTGGCCAGGAGCCAGCAAGATTGATTGATTGATTGATTGATTGATTGATTGATTGATTG 600
 QY 201 ArgArgAlaPheThrIysGlnSerProMetSerSerAlaArgGlnValSerLeuLeuIle 220
 Db 601 CGTAGAGCATTTCTAAGAGTGCCTTAGAGCTCGCAGACCAAGTTTCGCTTTTATA 660
 QY 221 LeuGlnSerPheLeuLeuMetProAspArgGlnIysGlnIysValIleGlnIysTyr 240
 Db 661 CTGAGCTTTTCTTTTAAATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 720
 QY 241 IleIysAspGlnIleGlnThrAlaIleValAlaValAlaTyrPargIysArgLeuMetThrGlnIys 260
 Db 721 ATTAAGATGAGGCGAGAGACGCTGCTGCTGCTGAGGAAAAGTTGATGACTAGAGA 780
 QY 261 GlnLeuAlaAlaIleGlnIysMetAspAlaArgGlnIysLeuLeuLeuValAlaCysPhe 280
 Db 781 GGATTGCTGGCGCTAGAAAATGATGCAAGGGGTTTCTTTTACTGATTCCTGTTT 840
 QY 281 GlnValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlnIys 300
 Db 841 GGTGTCCTTCAAACTTTAGAGTACAGATTGCTGATTTGATTAAGATGAGTGTG 900
 QY 301 AsnGlnIleAlaGlnIleLeuIysArgSerGlnPheLeuValProMetValSerGlnIle 320
 Db 901 AATGAGATTCCCGCTCTTGAAGCGGTCAAGATTCTTCTGCTCCATGCTTCAGGTATA 960
 QY 321 ValGlnSerSerIleIysArgGlnIysMetHisIleGlnAlaLeuGlnMetValTyrThrPhe 340
 Db 961 GTTGAAATCAAGTACAGCGTGAATGCAATTAAGCTCTTGAGATGGTATTATCTTT 1020
 QY 341 GlyMetGlnAspIysPheSerAlaAlaLeuValLeuThrSerPheLeuIysMetSerIys 360
 Db 1021 GGCATGAGGATTAAGTTTCAAGCTGCTAGATTCTAATCTTCAATTTTAAGATGACCAAG 1080
 QY 361 GlnSerPheGlnArgAlaIysArgIysAlaGlnSerProLeuAlaPheIysGlnIleAla 380
 Db 1081 GAGTCATTGAGGGGCAAAACGAAAGCCCACTCAACCTGCACTTAAAGAGCGGCT 1140
 QY 381 ThrIysGlnLeuAlaValIleuSerSerValMetGlnCysMetGlnThrHisIysLeuAsp 400
 Db 1141 ACAAGAGAGTACCTGTTATATCATCATGATTATGAGATGATGAGATGATGAGATGATG 1200
 QY 401 ProAlaIysGlnLeuProGlnIleIysGlnGlnIleValSerIleGlnIysAsp 420

Ds	1201	CCTCGAAGAACTACACAGGATGCCAGATCAAGAGCAAAATGTTAGCTTGGAGAAAGAC	1260			
Qy	421	ThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGluGlu	440			
Ds	1261	ACTCTTCAGCTCGCAAAAGAGATGGAAGAGAAAGCAAGATCTCTCAGTCTTAATGGAGAA	1320			
Qy	441	AlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSerProMet	460			
Ds	1321	GCACCACCTTGCACAGAAATGTATACCAACAGATAAAAGCTCCAGAGTTGTCAACCATG	1380			
Qy	461	GluMetProProValThrSerSerTyrSerProIleTyrArgAspArgSerPhePro	480			
Ds	1381	GAAATGCCACCAAGTAATCTTCTTCATCGTATCTCTTATCTACCGTGATGAAGCTTCTCT	1440			
Qy	481	SerGlnArgAspAspAspGlnAspGluIleSerAlaLeuValSerSerTyrLeuGlyPro	500			
Ds	1441	AGTCAAAAGACAGATGACCAAGATGAATATACAGCTCTTGAGTAGTATACCTCGGCCCG	1500			
Qy	501	SerThrSerPheProHisArgSerArgArgSerProGluTyrMetValProLeuProHis	520			
Ds	1501	TCAACATCTTCTCTCATCGCTCAAGAGATCCCGGAATATATGGTTCCCATTCACAT	1560			
Qy	521	GlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSerPro	540			
Ds	1561	GGTGGGTAGGAGAAAGTATATGTCATATGACATCTGGCCCAAAATTCATCTCTCCA	1620			
Qy	541	GlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArgHis	560			
Ds	1621	GGTCAAGGACATAGACTTCATCGAGTACTCTCGCTCTTGGTTTCAGACAGACAT	1680			
Qy	561	ProLeuGlnTyrSerProProIleHisGlyGlnGlnGlnLeuProTyrGlyIleGlnArg	580			
Ds	1681	CCACTACAGTACTCTCTCCCAATTCATGGACAAACACAGTATACCATATGATATACAAAG	1740			
Qy	581	ValTyrArgHisSerProSerGluGluArgTyrLeuGlyLeuSerAsnGlnArgSerPro	600			
Ds	1741	GTATTACAGACATTCACCATCTGAGAAAGATATTGGTTTATCCAAATCAAAGTCTCTCT	1800			
Qy	601	ArgSerAsnSerLeuAspProLys	609			
Ds	1801	CGCAGTAATCATCTTAGACCCCAAA	1827			
RESULT 2						
LOCUS	AX032763	2257 bp	DNA linear PAT 21-SEP-2000			
DEFINITION	Sequence 3 from Patent WO0046358.					
ACCESSION	AX032763					
VERSION	AX032763.1	GI:10279739				
KEYWORDS	synthetic construct					
SOURCE	synthetic construct					
ORGANISM	artificial sequences.					
REFERENCE	1					
AUTHORS	Dean,C., West,J. and Johanson,U.					
TITLE	Plant gene					
JOURNAL	Patent: WO 0046358-A 3 10-AUG-2000;					
FEATURES	DEAN CAROLINE (GB) ; WEST JOANNE (GB) ; PLANT BIOSCIENCE LTD (GB) ;					
	JOHANSON URBAN (SE)					
	Location/Qualifiers					
source	1..2257					
/organism="synthetic construct"						
/mol_type="unassigned DNA"						
/db_xref="taxon:32630"						
/note="likely cDNA sequence of the H51 FRI gene"						
ORIGIN						
Alignment Scores:						
Pred. No.:	3.64e-195	Length:	2257			
Score:	3104.00	Matches:	609			
Percent Similarity:	100.00%	Conservative:	0			
Best Local Similarity:	100.00%	Mismatches:	0			
Query Match:	100.00%	Indels:	0			
DB:	6	Gaps:	0			

US-09-890-475-1 (1-609) x AX032763 (1-2257)			
Qy	1	MetSerAsnTyrProProThrValAlaAlaGlnProThrThrThrAlaAsnProLeuLeu	20
Ds	362	ATGTCCAATTTATCCACCGAGCTGGCGGCAACCAACGAGCGGGAATCCCATGCTG	421
Qy	21	GlnArgHisGlnSerGluGlnArgArgGluLeuProLysIleValGluThrGluSer	40
Ds	422	CAGCGACATCAATCTGAACAGCGACGAGAGAAATACCGAAGATTGTGGAACAGAGTCT	481
Qy	41	ThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLysSerIleAsp	60
Ds	482	ACAAGTATGGACATTCAGCATCGTCAATCTAAGCAGCCTCAATTTTGAATCCATAGAC	541
Qy	61	GluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAsnAspLeuGln	80
Ds	542	GAAATAGCTGCGCTTTTCAGTTGCAGTGGAAACATTTCAACGCCAATTTGATGATCTT	601
Qy	81	LysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsnGlyValVal	100
Ds	602	AAGCACATCGAGTCAATCGMAAACGAATTTGATTCCTCAAACTCGAGAGTAACCGCGCT	661
Qy	101	LeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerProProArgAsnAsnVal	120
Ds	662	CTCGCGCGCGGAAACAATTTTCATCAGCCGATGTTATCCCTCCCGCGGAACAATGTA	721
Qy	121	SerValGluThrThrValThrValSerGlnProSerGlnGluIleValProGluThrSer	140
Ds	722	TCTGTAGAAACCAACCGCTCACTGTGAGCAACCGTCTCAGGAGATTGTACCGGAGACGTCG	781
Qy	141	AsnLysProGluGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyr	160
Ds	782	AATAAACCGGAGGGGAGCGTATGTGTGAGTTGATGTGTAGCAAAAGGTTCTCGGTAAATAC	841
Qy	161	IleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAlaLeuLys	180
Ds	842	ATATACCGCAATATCTCTGATCAAGTAAGTTAATGGAAGAGATTCTTCACGTTTGAAA	901
Qy	181	LeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGly	200
Ds	902	TTGGCCACAGGAGCCAGCGAAGTTTGTATTGGATTGTATTGGCAAGTTTACTTACAAGGG	961
Qy	201	ArgArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuIle	220
Ds	962	CGTAGAGCAATTTACTAAAGAGTCGCTATGAGCTCTGCGAGACAAGTTTCGCTTCTATA	1021
Qy	221	LeuGluSerPheLeuLeuMetProAspArgGlyLysGlyLysValLysIleGluSerTyr	240
Ds	1022	CTGGAGTCTTTCTTCTTAATGCTGATCGTGTGTAAGGGAAGGTGAAGATTGAGAGTTGG	1081
Qy	241	IleLysAspGluAlaGluThrAlaAlaValAlaAlaTyrArgLysArgLeuMetThrGluGly	260
Ds	1082	ATTAAGATGAGCGGAGACGCTGCTGTGTGTTGGAGGAAAGGTTGATGACTGAAGGA	1141
Qy	261	GlyLeuAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPhe	280
Ds	1142	GGATTAGCTGCGCTCAGAAAATGGATGCAAGGGGTTTCTCTTTACTAGTTGCTTTT	1201
Qy	281	GlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySer	300
Ds	1202	GGTGTCTCTTCAAACTTTTAGGAGTACAGATTTCGTGATTTGATAAGGATGAGTGTTCG	1261
Qy	301	AsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIle	320
Ds	1262	AATGAGATTGCGGCTGCTTTGAAGCGGTCAAGTTTCTTGTCCCTATGCTCAGGTATA	1321
Qy	321	ValGluSerSerIleLysArgGlyMetHisIleGluAlaLeuGluMetValTyrThrPhe	340
Ds	1322	GTTGAATCAAGTATCAAGCGTGGAAATGTCATATTGAAGCTCTTGAAGTGTATATACCTTT	1381
Qy	341	GlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLys	360

```
Db 1382 GGCATGGAGGATAAGTTTTCAGCTGCTCTAGTCTTAACCTTCAATCTTAAAGATGAGCAAG 1441
Qy 361 GluSerPheGluAArgAlaLysArgLysAlaGlnSerProLeuAlaPheLysGluAlaAla 380
Db 1442 GAGTCATTTGAGAGGCGAAACGGAACCCAGTCAGCTGGCATTTAAAGAGCGGCT 1501
Qy 381 ThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeuAsp 400
Db 1502 ACAAGCAGCTAGCTGTGTTATCATCATGTTATGAGTGTATGGAGACTCAAGTTAGAT 1561
Qy 401 ProAlaLysGluLeuProGlyTTPGlnLysGluGlnLeuValSerLeuGluLysAsp 420
Db 1562 CCTCGAAGAACATACAGGATGGCAGATCAAGAGCAATTTGTAGCTTGAGAAAGAC 1621
Qy 421 ThrLeuGlnLeuAspLysGluMetGluGlnLysAlaArgSerLeuSerLeuMetGluGlu 440
Db 1622 ACTCTTCAGCTCGCAAAAGAGATGGAAGAGAAAGCAAGATCTCTCAGTTTAAATGGAGAA 1681
Qy 441 AlaAlaLeuAlaLysArgMetTyrAsnGlnGlnLysArgProArgLeuSerProMet 460
Db 1682 GCGGCACTTGGCAAGAGATGTATTAACCAACAGATTAACCTCCAGGTTGTCAACCATG 1741
Qy 461 GluMetProProValThrSerSerSerTyrSerProLysTyrArgSerPhePro 480
Db 1742 GAAATGCCACCAAGTAATCTTTCATCGTATTTCTCTATCTACCGGTGATAGAAGCTTTCT 1801
Qy 481 SerGlnArgAspAspGlnAspGluLysSerAlaLeuValSerSerTyrLeuGlyPro 500
Db 1802 AGTCAAGAGACGATGACCAAGATGAATATCAGCTCTTGTAGTAGTACTCTGGCCG 1861
Qy 501 SerThrSerPheProHisArgSerArgSerProGluTyrMetValProLeuProHis 520
Db 1862 TCAACATCTTTTCTCATCGCTCAAGAGATGCCCGGAATATATGGTTCCACTTCCACAT 1921
Qy 521 GlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSerPro 540
Db 1922 GGTGGGTAGGAAGAGGTATATGATATGATGATGATGATGATGATGATGATGATGATGAT 1981
Qy 541 GlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArgHis 560
Db 1982 GGTCAAGGACATAGACTTCATCGACAGTACTCTCGCTCTTTGGTTTCAACGACAGACAT 2041
Qy 561 ProLeuGlnTyrSerProProLysGlyGlnGlnGlnLeuProTyrGlyLeuArg 580
Db 2042 CCACTACAGTACTCTCTCTCAATCATGGACAAACAGTACCATATGGTATACAAAGG 2101
Qy 581 ValTyrArgHisSerProSerGluArgTyrLeuGlyLeuSerAsnGlnArgSerPro 600
Db 2102 GTTACAGACATTCACCATCTGAAGAGATATTTGGGTTTATCCAAATCAAGGTCTCT 2161
Qy 601 ArgSerAsnSerSerLeuAspProLys 609
Db 2162 CGCAGTAATCATCATATAGACCCCAAA 2188

RESULT 3
AX032762
LOCUS AX032762
DEFINITION Sequence 2 from Patent WO046358.
ACCESSION AX032762
VERSION AX032762.1 GI:10279738
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
AUTHORS Dean,C., West,J. and Johanson,U.
TITLE Plant gene
JOURNAL Patent: WO 0046358-A 2 10-AUG-2000;
DEAN CAROLINE (GB) ; WEST JOANNE (GB) ; PLANT BIOSCIENCE LTD (GB) ;
JOHANSON URBAN (SE)
FEATURES
Location/Qualifiers
```

```
source 1..3761
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Alignment Scores:
Pred. No.: 9,02e-187 Length: 3761
Score: 2881.00 Matches: 608
Percent Similarity: 78.96% Conservatives: 0
Best Local Similarity: 78.96% Mismatches: 1
Query Match: 96.04% Indels: 162
DB: 6 Gaps: 2

US-09-890-475-1 (1-609) x AX032762 (1-3761)
Qy 1 MetSerAsnTyrProProThrValAlaAlaGlnProThrThrThrAlaAsnProLeuLeu 20
Db 574 ATGTCCCAATATATCCACCGAGCGTGGCGCAACCCACACACGCGCAATCCATGCTGTG 633
Qy 21 GlnArgHisGlnSerGluGlnArgArgArgGluLeuProLysLeuValGluThrGluSer 40
Db 634 CAGCGACATCATCTGACACGCGACGAGAGAAATACCGAAGATTGTGAAACAGAGTCT 693
Qy 41 ThrSerMetAspLysThrLeuGlnSerLysGlnProGlnPheLeuLysSerLeuAsp 60
Db 694 ACAAGTATGGACATATACGATCGTCAATCTAAGCAGCTCAATTTTGAATCCATAGAC 753
Qy 61 GluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAspAspLeuGln 80
Db 754 GAATAGCTGCGTTTTCAGTTGAGTGGAAACATTCAAACGCAATTCGATGATCTTCAG 813
Qy 81 LysHisLeuGluSerLeuGluAsnAlaLysSerLysLeuGluSerLeuValVal 100
Db 814 AAGCACATCGAGTCAATCGAAACGCAATTCATTCAAACTCGAGAGTAACGCGCTGTGTC 873
Qy 101 LeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerProProArgAsnVal 120
Db 874 CTCGCGCGCGGAAACAAATATTTCCATCAGCGCAATGTTATCGCTCCGCGGAAACATGTA 933
Qy 121 SerValGluThrThrValThrValSerGlnProSerGlnGluLeuValProGluThrSer 140
Db 934 TCTGTAGAACCAACCGCTCACTGTGAGCCAAACCGTCTCAGAGATTGTACCGAGACGTG 993
Qy 141 AsnLysProGluGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyr 160
Db 994 AATAAACCGGAGGCGGAGCTATGCTGTGATGTTGATGTAGCAAGGCTCTGCGTAAATAC 1053
Qy 161 IleTyrAlaAlaAsnLysSerAspGlnAlaLysLeuMetGluGluLeuProSerAlaLeuLys 180
Db 1054 ATATACCGAATATCTCTGATCAAGCTAAGTTAATGGAAGAGATTCTTCAGCTTGAAA 1113
Qy 181 LeuAlaLysGluProAlaLysPheValLeuAspCysLysGlyLysValLysLeuGlnGly 200
Db 1114 TTGGCCAAAGGAGCGAGCAAGTTTGTATGATTTGATTTGGCAAGTTTACTTACAGGG 1173
Qy 201 ArgArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeu 220
Db 1174 CGTAGACATTTACTTAAGAGTCCCTATGAGCTCTGCGAGACAAAGTTTCGTTCTTATA 1233
Qy 221 LeuGluSerPheLeuLeuMetProAspArgGlyLysGlyLysValLysLeuSerTyr 240
Db 1234 CTGGAGTCTTTTCTTCTAATGCTGATCGTGTGTAAGGGAAGGTGAAGATTGAGAGTTGG 1293
Qy 241 IleLysAspGluAlaGluThrAlaAlaValAlaTyrArgLysArgLeuMetThrGluGly 260
Db 1294 ATTAAGATGAGGCGGAGACGCTGCTGTGCTGGAGGAAAGGTTGATGACTGAGGA 1353
Qy 261 GlyLeuAlaAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPhe 280
Db 1354 GGATTACCTCGGCTGAGAAATGGATGCAAGGGTTTGCTTTTACTAGTTGCTGTTT 1413
Qy 281 GlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuLeuMetSerGlySer 300
```


Db 1414 GGTGTTCTTCAAACTTTAGGAGTACAGATTGCTGATTTGATTAAGGATGAGTGGTTCC 1473
Qy 301 AsnGluLeuAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSer----- 318
Db 1474 AATGAGATTGCGGTGCTTTGAAGCGGTACAGATTCTCTTGTCCTATGGTCTCAGGTACC 1533
Qy 318 ----- 318
Db 1534 ATATTCTGTTCTCACTCGGTGAATTTCATTCAAAGGTGGTTCCTTTGTTGATCATCATC 1593
Qy 318 ----- 318
Db 1594 GACCAACATCAAGTTCACATCTTTGTTTTCGATAAGCTTGATGATAAATAGGAGAC 1653
Qy 318 ----- 318
Db 1654 ACATCAATATTAGAGTGCAATGACTGATTCAGTTCAGCAAAATCCTAGTAGAAATTAATCTG 1713
Qy 318 ----- 318
Db 1714 GAAAGAACTTGGAACTCTCAACCATAGGTTTGGTACGAAATGTTGTTGTCAGAACCA 1773
Qy 318 ----- 318
Db 1774 AATGATAGGCTATTGCCTTGAATAGTGTCTTGTGTTCCATATATGGAACCTTAAA 1833
Qy 318 ----- 318
Db 1834 TCGTATGACTTAGCTGTTGGATACCTAATAAGCTTAAGCAATGCCAATCTTAAGAGTGG 1893
Qy 319 -----GlyileValGluSerSerIleLysArgGlyMet 329
Db 1894 TACTTACAAATATCTATTGTCATAGTATAGTTGATCAAGTATCAAGCTGGATG 1953
Qy 330 HisLeuAlaLeuGluMetValThrPheGlyMetGluAspLysPheSerAlaAla 349
Db 1954 CATATTGAAGCTCTTGAGATGTTTATACCTTTGGCATGGAGGATAAGTTTTCAGCTGCT 2013
Qy 350 LeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLys 369
Db 2014 CTAGTTCTTAATCTTCTTAAGATGAGCAAGGAGTCATTTGAGAGGGCAAAACGGAAA 2073
Qy 370 AlaGlnSerProLeuAlaPheLys----- 377
Db 2074 GCCAGTACCGCTGGCATTGTTGT-ATGAACCTTCCCTTCACATATGATACCTTTATGA 2132
Qy 378 -----GluAla 379
Db 2133 ACTCTTTATCATCTCTGAGTCTGACCATTCATATATTTATTTCTCAACAGAAAGAGCG 2192
Qy 380 AlaThrLysGlnLeuAlaValLeuSerValMetGlnCysMetGluThrHisLysLeu 399
Db 2193 GCTACAAAGCAGCTAGCTGTGTTATCATCAGTTATGCACTGATGAGACTCACAAGTTA 2252
Qy 400 AspProAlaLysGluLeuProGlyTrpGlnLysGluGlnLysValSerLeuGluLys 419
Db 2253 GATCTGCGAAAGAACTACCAAGGATGGCAGATCAAGAGCAAAATTTGTTAGCTTCGAGAAA 2312
Qy 420 AspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGlu 439
Db 2313 GACACTCTTCAGCTCGACAAAGAGATGGAAGAGAAAGCAAGATCTCTCAGTTAATGAG 2372
Qy 440 GluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnLysArgProArgLeuSerPro 459
Db 2373 GAAAGCGCACTTCCCAAGAGATGTATAACCAACAGATAAAACGTCGAAGTGTGTCAACC 2432
Qy 460 MetGluMetProProValThrSerSerSerTyrSerProIleTyrArgAspArgSerPhe 479
Db 2433 ATGGAATATGCCACAGTAACTTCTTCATCGTATTCCTCTACTACTACCGGTATAGAGCTTT 2492
Qy 480 ProSerGlnArgAspAspGlnAspGlnLeuSerAlaLeuValSerSerTyrLeuGly 499

Db 2493 CCTAGTCAAAGAGACGATGACCAAGATGAATATCAGCTCTTGTGAGTAGTACTCTCGC 2552
Qy 500 ProSerThrSerPheProHisArgSerArgSerProGluTyrMetValProLeuPro 519
Db 2553 CGGTCAACATCTTTTCTCATCGCTCAAGAAGATCCCGGAATATATGTTCCACTTCCA 2612
Qy 520 HisGlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSer 539
Db 2613 CATGTGGGTAGGAAGATGTATGATATGAACATCTGCCCCCAAAATTCATACTCT 2672
Qy 540 ProGlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArg 559
Db 2673 CAGGTCAGGACATAGACTTCATCGACAGTACTCTCGTCTTTGTTCCAGACAGAGA 2732
Qy 560 HisProLeuGlnTyrSerProPheHisGlyGlnGlnLeuProTyrGlyGln 579
Db 2733 CATCCACTACGACTCTCTCTCAATTCATGACAAACACAGATTACCATATGTTATACAA 2792
Qy 580 ArgValTyrArgHisSerProSerGluGluArgTyrLeuGlyLysSerAsnGlnArgSer 599
Db 2793 AGGTTTACAGACATTCACATCTCAAGNAAGATATTGGGTATTATCCAATCAAGGCT 2852
Qy 600 ProArgSerAsnSerSerLeuAspProLys 609
Db 2853 CCTCGAGTAACTCATCTAGACCCCAA 2882

RESULT 4
AP228499
LOCUS AP228499 3761 bp DNA linear PLN 14-OCT-2000
DEFINITION Arabidopsis thaliana FRIGIDA (FRI) Gene, complete cds.
ACCESSION AP228499
VERSION AF228499.1 GI:10801173
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 3761)
AUTHORS Johanson,U., West,J., Lister,C., Michaels,S., Amasino,R. and Dean,C.
TITLE Molecular analysis of FRIGIDA, a major determinant of natural variation in Arabidopsis flowering time
JOURNAL Science 290 (5490), 344-347 (2000)
MEDLINE 20485641
PUBMED 11030654
REFERENCE 2 (bases 1 to 3761)
AUTHORS Johanson,U. and Lister,C.
TITLE Direct Submission
JOURNAL Submitted (25-JAN-2000) Molecular Genetics, John Innes Centre, Colney Lane, Norwich, Norfolk NR4 7UH, U.K.
FEATURES
Location/Qualifiers
1..3761
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="ecotype H51"
/db_xref="taxon:3702"
/chromosome="4"
/map="between ml51 and ml22"
/gene="FRI"
/gene="FRI"
join(<574..1528,1922..2094,2184..>2885)
/product="FRIGIDA"
/gene="FRI"
join(574..1528,1922..2094,2184..2885)
/notes="late flowering gene"
/codon_start=1
/product="FRIGIDA"
/protein_id="AG23414.1"
/db_xref="GI:10801174"
/translation="MSNIPTVAQPTTTANPILQHQSEQRRELPKIVETESTSMD
ITIGSKQFQFLKSIDELAASFVAVETFRKQFDDLDQKHIESINLDSKLESGVVIA

ARNNHFPMLSPNNVSVETTVTSQSOEIVPETSNTKPEGRMCMLCSKGLRKY
 IYANISDAKMBEIPSAKIAKBPFAKFDICIKFYLOGRAFTKESPMSSAROVSL
 LILNIFLMPDRGKVKYIESWIKDEAETAFAVAKRLMTGEGGIAAEKMDARGLIL
 VACFVPSNFEDTDLIDILMSGNEIAGALKRSQPIYPMVSGIVSSIKRGWHEIAL
 EMVYTFGMEFSAALVUTISFLKMSKSFERAKRAQSPFLAKFAATKQYALVSSWQ
 CMETHKLDPAKELPQWIKQIVLEKDTLQDKEMEEKAKSLSLMEERAAKAMWQ
 QIKRPLSPMBMPVTSSSYSPIDYRDSFPPSDGDDQDETSALVSSYGPSTSPHRS
 RRSPPYVPLPHGGIGRVSAYAYELHAPNSYSPGHGRLHRYQSPSLVHGQHPLOYSP
 PTHGQQQLPYGIQRYVRHSPSEERYLGLNSRSPNSSLDPK"

ORIGIN

Alignment Scores:

Pred. No.: 9,026-187 Length: 3761
 Score: 2981.00 Matches: 608
 Percent Similarity: 78.96% Conservativity: 0
 Best Local Similarity: 78.96% Mismatches: 1
 Query Match: 96.04% Indels: 162
 DB: 8 Gaps: 2

US-09-890-475-1 (1-609) x AF228499 (1-3761)

Qy 1 MetSerAsnTyrProProThrValAlaAlaGlnProThrThrAlaAsnProLeuLeu 20
 Db 574 ATGTCCTCAATATCCACCGACGGTGGCGGCGCAACCAACACGCGCAATCCACTGCTG 633
 Qy 21 GlnArgHisGlnSerGluGlnArgArgGluLeuProLysIleValGluThrGluSer 40
 Db 634 CAGCGCATCAATCTCTGAACAGCGACGAGAGAAATACCGAAGATTGTCCAAACAGAGTCT 693
 Qy 41 ThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLysSerIleAsp 60
 Db 694 ACAAGTATGACATTAACGATCGGTCAATCTAAGCAGCCTCAATTTTGAATCCATAGAC 753
 Qy 61 GluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAspSerLeuGln 80
 Db 754 GAATTAGCTCGCTTTTTCAGTTCAGTGGAAACATTCAACGCGCAATTCGATATCTTCAG 813
 Qy 81 LysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsnGlyValVal 100
 Db 814 AAGCACATCGATCAATCGAAGCAATGATTCCAACTCGAGATTAACGCGTGTGC 873
 Qy 101 LeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerProProArgAsnAsnVal 120
 Db 874 CTCGCGCGCGGAACAATAATTTCCATCAGCGCATGTTATCGCTCCGCGGAACAATGTA 933
 Qy 121 SerValGluThrValThrValSerGlnProSerGlnGluIleValProGluThrSer 140
 Db 934 TCTGTAGAACACCCCTACTGTGAGCCCAACCGTCTCAGAGATTTGACCGAGAGCTCG 993
 Qy 141 AsnLysProGluGlyGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyr 160
 Db 994 AATAAACCGGAGGGGAGCGTATGTGTGAGTTGATGTAGCAAGGTCTGCGTAAATAC 1053
 Qy 161 IleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAlaLeuLys 180
 Db 1054 ATATACGCGAATATCTCTGATCAAGTATAGTTAATGGAAGAGATTCCCTTCAGCTTTGAAA 1113
 Qy 181 LeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGly 200
 Db 1114 TTGCGCCAGAGAGCCAGCAAGTTTGTATTGATTGTTGCAAGTTTACTTACAAGGG 1173
 Qy 201 ArgArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuIle 220
 Db 1174 CGTAGAGCATTTACTAAAGAGTCGCTATGAGTCTGCGAGACAAGTTTCGCTTCTTATA 1233
 Qy 221 LeuGluSerPheLeuLeuMetProAspArgGlyLysGlyLysValLysIleGluSerTrp 240
 Db 1234 CTGGAGTCTTTTCTTAATGCTGTATCGTGGTAAAGGGAAGGTGAAGATTGAGATTGG 1293
 Qy 241 IleLysAspGluAlaGluThrAlaAlaValAlaTrpArgLysArgLeuMetThrGluGly 260
 Db 1294 ATTAAGATGAGCGGAGCGGTGCTGTGCTTGGAGGAAAAGGTTGATGACTGAAGGA 1353

Qy 261 GlyLeuAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPhe 280
 Db 1354 GGATTAGCTGCGGCTGAGAAAAATGGATCAAGGGGTTCCTTTTACTAGTTCCTTGT 1413
 Qy 281 GlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySer 300
 Db 1414 GGTGTTCCCTTCAACTTTAGGAGTACAGATTTCGTGGATTTCATAGGATGAGTGGTTCG 1473
 Qy 301 AsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSer 318
 Db 1474 AATGAGATTGCGGCTGCTTTGAAGCGGTCACAGTTTCTTGTCCCTATGCTCTCAGGTACC 1533
 Qy 318 ----- 318
 Db 1534 ATATTCTGTTCTCAGTCGGTGAATTTTCATTGCAAGAGTGGTTCCTTTTCTTGACATCATC 1593
 Qy 318 ----- 318
 Db 1594 GACCAACATCAAGTTCCTCTTTTTCGATAGCTTTGATGGTATATAAACTAGGAGAGC 1653
 Qy 318 ----- 318
 Db 1654 ACATCAAAATATTTAGAGTGCATGACTGATTGAGCCAAATCCTAGCTAGAAATTAATCTG 1713
 Qy 318 ----- 318
 Db 1714 GAAAGAACTTGGAACTCTCAACCATAGTTCCTTGGTACGAAATGTTGCTTGTTCAGAACCA 1773
 Qy 318 ----- 318
 Db 1774 AATGATAGGCTATTGCGCTTGAANAATAGTTCCTTGTGGTTCCTCAATATTGGAAGTTAAAA 1833
 Qy 318 ----- 318
 Db 1834 TCGTATGACTTAGCTGTGGTACTAATTAAGCTTAAGCAATGCCAACTCTAAGAAAGTGG 1893
 Qy 319 ----- GlyIleValGluSerSerIleLysArgGlyMet 329
 Db 1894 TACTTACACAATATTCCTATTGTCATAGGTATAGTTGAATCAAGTATCAAGCGTGAATG 1953
 Qy 330 HistIleGluAlaLeuGluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAla 349
 Db 1954 CATATTGAAGCTCTTGAGATGGTTTATACCTTTGGCATGGAGGATAAGTTTTCAGCTGCT 2013
 Qy 350 LeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLys 369
 Db 2014 CTAGTTCCTAACTTCATCTTTAAAGATGAGCAGAGGTCATTTTGAGAGGGCAAAACGAAA 2073
 Qy 370 AlaGlnSerProLeuAlaPheLys----- 377
 Db 2074 GCCCAGTCACCGCTGGCATTTGT-ATGAACCCCTTCCCTTGCACATTATGTACCTTTATGA 2132
 Qy 378 ----- GluAla 379
 Db 2133 ACTCTTTATCATCATCTGAGTCTGACCATTTGATATATTTTCTCAACAGAAAGAGCG 2192
 Qy 380 AlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeu 399
 Db 2193 GCTCAAGAGCAGCTAGCTGTGTTATCATCATGATTGATGAGTGTATGAGACTCACAGTTA 2252
 Qy 400 AspProAlaLysGluLeuProGlyTyrGlnIleLysGluGlnIleValSerLeuGluLys 419
 Db 2253 GATCTCGGAAAGAACTACCAAGGATGCGAGATCAAGAGCAAAATTTGTAGCTTGGAGAAA 2312
 Qy 420 AspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGlu 439
 Db 2313 GACACTCTTCAGCTCGCAAAAGAGATGGAAGAGAAACAGATCTCTCAGTTTATGAG 2372
 Qy 440 GluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSerPro 459
 Db 2373 GAAGCGCAGCTTGCACAGAGATGATAAACCAACAGATAAAACGTCCTCAAGGTTTCTCACC 2432
 Qy 460 MetGluMetProProValThrSerSerTyrSerProIleTyrArgAspArgSerPhe 479


```
Db 1114 ATTAAGATGAGCGGAGAGCGCTCTGTTGCTTGGAGAAAGGTTGATGACTGAGGA 1173
Qy 261 GlyLeuAlaAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPhe 280
Db 1174 GGATTAGCTCGGCTGAGAAAATGGATGCAAGGGTTTGCTTTTACTAGTTGCTGTGTTT 1233
Qy 281 GlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuLeuArgMetSerGlySer 300
Db 1234 GGTGTTCTCTCAACATTTAGGAGTACAGATTTGCTGGATTTGATGAAGATGATGGTTG 1293
Qy 301 AsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSer 318
Db 1294 AATGAGATTGCCGGTGCTTTGAAGCGGTACAGTTTCTTGCCCTATGGTCTCAGGTACC 1353
Qy 318 ----- 318
Db 1354 ATATTCTGTTCTCACTCGGTGAATTTCAATGCAAAAGTGGTTCCTTTTGTGACATCATC 1413
Qy 318 ----- 318
Db 1414 GACCAACATCAAGTCCATCTTTGTTTTCGATAAGCTTGTATGATTAACCTAGGAGAGC 1473
Qy 318 ----- 318
Db 1474 ACATCAAAATATTAGAGTGCAATGACTGATTGAGCCAAATCCTAGCTAGAAATTAATCTG 1533
Qy 318 ----- 318
Db 1534 GAAAGAACTTGGAACTCTCAACCATAGGTTTGTGTAAGAAATGTTGTTGTGTCAGAACCA 1593
Qy 318 ----- 318
Db 1594 AATGATAGGCTATTGCCCTTGAATAAGTAGTGTCTTGTGTTTCCAAATATTGGAAGTTAAA 1653
Qy 318 ----- 318
Db 1654 TCATATGACTTAGCTGTGGATACTTAATTAAGCTTAAGCAATGCCCACTCTAAGAGTGG 1713
Qy 319 -----GlyIleValGluSerSerIleLysArgGlyMet 329
Db 1714 TACTTACACAATATTCTATTGTTGTCATAGGTATAGTTGAATCAAGTATCAAGCGTGAATG 1773
Qy 330 HisIleGluAlaLeuGluMetValThrPheGlyMetGluAspLysPheSerAlaAla 349
Db 1774 CATATTGAAGCTCTTGAGATGGTTTATACCTTTGGCATGGAGATGAAGTTTTCAGCTGCT 1833
Qy 350 LeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLys 369
Db 1834 CTAGTTCTTAACCTTCACTTTAAAGATGAGCAAGGAGTCAATTTGAGAGGCGCAAAACGGAAA 1893
Qy 370 AlaGlnSerProLeuAlaPheLys ----- 377
Db 1894 GCCAGTCACCGCTGGCATTTGT-ATGAACCTTCCTCTGACATTAATGATGACCTTTATGA 1952
Qy 378 -----GluAla 379
Db 1953 ACTCTTTTATCATCTGAGTCTGACCATTGATATATTATTCTCAACAGAGAAGAGCG 2012
Qy 380 AlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeu 399
Db 2013 GCTACAAAGCAGCTAGCTGTGTTATCATCAGTTATGCTGATGATGAGAGTATGAGAGTTA 2072
Qy 400 AspProAlaLysGluLeuProGlyTrpGlnIleLysGluGlnIleValSerLeuGluLys 419
Db 2073 GATCTCTCGAAGAACTACCAAGATGCGAGATCAAGAGCAAAATGTTAGCTTGGAGAAA 2132
Qy 420 AspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGlu 439
Db 2133 GACACTCTTCAGCTCCACAAAGAGATGGAAGAGAAAGCAAGATCTCTCAGTTTAAATGGAG 2192
Qy 440 GluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSerPro 459
```

```
Db 2193 GAAGCGCAGCTTCCCAAGAGATGTATATAACCAACAGATAAAACGTCCAAGGTTGTCAACC 2252
Qy 460 MetGluMetProValThrSerSerSerTyrSerProIleTyrArgAspArgSerPhe 479
Db 2253 ATGGAATATGCCACACAGTAATCTTTCATCGTATCTCTTATCTCCATCTACCGTGATGAAGCTTT 2312
Qy 480 ProSerGlnArgAspAspGlnAspGluIleSerAlaLeuValSerSerTyrLeuGly 499
Db 2313 CCTAGTCAAGAGACCATGACCAAGATGAATATATCAGCTCTTGTGAGTAGTTACCTCGGC 2372
Qy 500 ProSerThrSerPheProHisArgSerArgArgSerProGluTyrMetValProLeuPro 519
Db 2373 CCCTCAACATCTTCTCTCATCGCTCAAGAAGATCCCGGAATATATGTTCCACTTCCA 2432
Qy 520 HisGlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSer 539
Db 2433 CATGGTGGTTAGGAAGAGTGTATATGCATATGAACATCTGCCCCCAAAATTCATACTCT 2492
Qy 540 ProGlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArg 559
Db 2493 CCAGGTCAACGACATAGACTTCATCGACAGTACTCTCCGTCTTGGTTCCACGACAGAGA 2552
Qy 560 HisProLeuGlnTyrSerProProIleHisGlyGlnGlnGlnLeuProTyrGlyIleGln 579
Db 2553 CATCCACTACGACTCTCTCTCAATTCATGCAACAACAGTACCATATGGTATACAA 2612
Qy 580 ArgValTyrArgHisSerProSerGluArgTyrLeuGlyLeuSerAsnGlnArgSer 599
Db 2613 AGGGTTTACAGACATTCACCATCTGAAGAAGATATTTGGGTTTATCCCAATCAAGGTCT 2672
Qy 600 ProArgSerAsnSerSerLeuAspProLys 609
Db 2673 CCTCGCAGTAACATCATTTAGACCCAAA 2702

RESULT 6
AV198402
LOCUS AV198402 3261 bp DNA linear PLN 12-JUN-2003
DEFINITION Arabidopsis thaliana ecotype KZ-9 FRIGIDA (FRI) gene, complete cds.
ACCESSION AV198402
VERSION AV198402.1 GI:31558912
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsia.
REFERENCE
1 (bases 1 to 3261)
Gazzani,S., Gendall,A.R., Lister,C. and Dean,C.
Analysis of the molecular basis of flowering time variation in
Arabidopsis accessions
Plant Physiol. 132 (2), 1107-1114 (2003)
2 (bases 1 to 3261)
Gazzani,S. and Dean,C.
Direct Submission
Submitted (17-DEC-2002) Cell and Developmental Biology, John Innes
Center, Colney Lane, Norwich NR4 7UH, UK
LOCATION/Qualifiers
FEATURES
1..3261
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="4"
/map="between mi51 and mi122"
/ecotype="KZ-9"
<505..>2816
/gene="FRI"
join(<505..1459,1853..2025,2115..>2816)
/product="FRIGIDA"
join(505..1459,1853..2025,2115..2816)
/gene="FRI"
/note="flowering gene"
/codon_start=1

gene
mRNA
CDS
```

```
/product="PRIGIDA"  
/protein_id="AAP49808.1"  
/db_xref="GI:11558913"  
/translation="MSNYPFVVAQAQPTTANPLLRHQSQRRRLPKIVETSTSD  
ITIGSKQPOLKSIDELAFSAVETFKQEDDLQKHIESIENADISKLESNGVIA  
ARNNFHQPLSPFRNNVSVETTVSQSQEIYPTSNRPEGGRMELMCSGLRKY  
IYANISQAKLMEIPKALAKAPEFVLDICIGFYLOGRRAFTKSPMSARQVSL  
LILESFLIMPDRGKVKIESWKDEAETAAVWRKELMTTEGGIAAAKMDARGLLIL  
VACFGVSPNRESTDLLIRMSGNEJAGALKRSQFLVPMVSGIVGESIKRGVHIBAL  
EMVTFQMEDKFSALVLTSPKMSKESFERAKKAQSPAPKANAQKQAVLSSVMQ  
CMETHKLDPAKELFGWQIKQIVSLEKDTQLDKEMEKEKASLSLSMEALAKRMNQ  
QIKRPLSPMPGPPVTSSVSYPIYRDSFPQDDQDEISALVSSVLTGFSPTFPRS  
RRSPYVPLPGLGHSVYAYEHLAPNSPSYGHGHLRQYSPSLVHGHRHPLQVSP  
PIHQQQQLPYGIQRYVHRSPSEERYLGLSNQSPRSNSLDPK"
```

ORIGIN

```
Alignment Scores:  
Pred. No.: 1,866-186 Length: 3261  
Score: 2975.00 Matches: 607  
Percent Similarity: 78.83% Conservative: 0  
Best Local Similarity: 78.83% Mismatches: 2  
Query Match: 95.84% Indels: 162  
DB: 8 Gaps: 2
```

US-09-890-475-1 (1-609) x AY198402 (1-3261)

```
Qy 1 MetSerAsnThrProThrValAlaAlaGlnProThrThrAlaAsnProLeuLeu 20  
Db 505 ATGTCCAATTATCCACCACGCGTGGCGCGCAACCCACACGACGCGCAATCCACTGCTG 564  
Qy 21 GlnArgHisGlnSerGluGlnArgArgGluLeuProLysIleValGluThrGluSer 40  
Db 565 CAGGACATCAATCTGAACACGCGACGACGAGAAATTCGGAAGATGTGCGAACAGAGTCT 624  
Qy 41 ThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLysSerIleAsp 60  
Db 625 ACAAGTATGGACATTACGATCGGTCAATCTAAGCAGCGCTCAAAATTTGAAATCCATAGAC 684  
Qy 61 GluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAspAspLeuGln 80  
Db 685 GAATTAGTCGGTTTCAGTTGCGATGGGAACATTCAAACGCCCAATTCGATGATCTTCAG 744  
Qy 81 LysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsnGlyValVal 100  
Db 745 AAGCACATCGAGTCAATCGAAACGCAATTGATTCCAAACCTCGAGAGTAACGCGGTGTC 804  
Qy 101 LeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerProProArgAsnAsnVal 120  
Db 805 CTCGCCGCGCGAAACAATAATTTCCATCAGCCGATGTTATCGCTCCGCGGGAACAATGTA 864  
Qy 121 SerValGluThrThrValThrValSerGlnProSerGlnGluIleValProGluThrSer 140  
Db 865 TCTGTAGAAACCAACCGTCACTGTGAGCCACCGTCTCAGGAGATTGTACCGGAGACGCTG 924  
Qy 141 AsnLysProGluGlyGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyr 160  
Db 925 AATTAACCGCGGGGGGACGATGTGTGTGAGTGTGATGTGTAGCAAGGCTGCGTAAATAC 984  
Qy 161 IleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAlaLeuLys 180  
Db 985 ATATACGCCAATATCTCTGATCAAGCTAAGTTATGAGAGATTCCTTCAGCTTTGAAA 1044  
Qy 181 LeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGly 200  
Db 1045 TTGGCCAAAGGACCGAGCAAGTTTGTATTGGAATTGATTGCAAGTTTACTTACAAGGG 1104  
Qy 201 ArgArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuIle 220  
Db 1105 CGTAGACATTACTTAAAGATCGCCTATGAGCTCTCGGACAGATTTGCTTCCTTATA 1164  
Qy 221 LeuGluSerPheLeuLeuMetProAspArgGlyLysGlyLysValLysIleGluSerTyr 240  
Db 1165 CTGGAGTCTTTTCTTAATGCTGATCGTGGTAAAGGAGGTTGAAGATTGACAGTTGG 1224
```

```
Qy 241 IleLysAspGluAlaGluThrAlaAlaValAlaIleThrAlaGlyLysArgLeuMetThrGluGly 260  
Db 1225 ATTAAGATGAGCGGAGAGCGCTGCTGTTGTTGGAGGAAAGGTTGATGACTGAGGA 1284  
Qy 261 GlyLeuAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuLeuValAlaCysPhe 280  
Db 1285 GGATTAGCTGCGGCTGAGAAAATGGAATGATGCAAGGGGTTGCTTTTACTAGTTCTGTTT 1344  
Qy 281 GlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySer 300  
Db 1345 GGTGTTCTCTCAACCTTAGAGAGTACAGATTGCTGGATTGATAGAGTAGTGGTTCG 1404  
Qy 301 AsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSer 318  
Db 1405 AATGAGATTGCGGCTGCTTTGAAGCGGTCAAGTTTCTTGTCCTATGCTCAGGTACC 1464  
Qy 318 ----- 318  
Db 1465 ATATTCGTCTTCACCTCGGTGAATTCATTGCAAAAGGTGGTTCCTTTGTTGACATCATC 1524  
Qy 318 ----- 318  
Db 1525 GACCAACATCAAGTTCATCTTTGTTTTCGATAAGCTTGATGTTATAAACTAGGAGAGC 1584  
Qy 318 ----- 318  
Db 1585 ACATCAAAATATTAGAGTGCATGACTGATTGAGCAAAATCCTAGCTAGAAATTAATCTG 1644  
Qy 318 ----- 318  
Db 1645 GAAAGAACTTGGAACTCTCAACCATAGGTTTTGGTAGCAAAATTTGTTGTTGTCAGAACCA 1704  
Qy 318 ----- 318  
Db 1705 AATGATAGGCTATTGCTTTGAAATAGTGTCTTCTGTGGTTTCCAATATTGGAAGTTAAA 1764  
Qy 318 ----- 318  
Db 1765 TCATATGACTTAGCTTTGGATACTAATTAAGCTTAAGCAATGCCAATCTTAAGAAGTGG 1824  
Qy 319 ----- GlyIleValGluSerSerIleLysArgGlyMet 329  
Db 1825 TACTTACACATATTCTATTGTCATAGGTATAGTTGAATCAAGTATCAAGCGTGAATG 1884  
Qy 330 HisIleGluAlaLeuGluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAla 349  
Db 1885 CATATTGAAGCTCTTGAGATGGTTTATACCTTTGGCATGGAGATAAGTTTTCAGCTGCT 1944  
Qy 350 LeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGluAlaLysArgLys 369  
Db 1945 CTAGTTCTTAATCTTCAATCTTAAAGATGAGCAAGGAGTCATTTGAGAGGGCAAAACGAAA 2004  
Qy 370 AlaGlnSerProLeuAlaPheLys ----- 377  
Db 2005 GCCCAGTCCCGCTGCAATTTGT-ATGAACCCCTCCCTTGCACATTATGTACCTTTATGA 2063  
Qy 378 -----GluAla 379  
Db 2064 ACTCTTTATCATCATCTGAGTCTGACCATTTGATATATTATTCTCAACAGAAAGAGCG 2123  
Qy 380 AlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeu 399  
Db 2124 GCTACAAACAGCTAGCTGTGTATCATCAGTATGCGATGATGAGACTCACAAGTTA 2183  
Qy 400 AspProAlaLysGluLeuProGlyTyrGlnIleLysGluIleValSerLeuGluLys 419  
Db 2184 GATCTCTGCGAAGAACTACCAAGTGGCAGATCAAGAGCAAAATTTGTAGCTTGGAGAAA 2243  
Qy 420 AspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGlu 439  
Db 2244 GACACCTCTTCAGTCCGACAAAGAGATGGAAGAAAGCAAGATCTCTCAGTTTAATGGAG 2303
```

```
QY 440 GluAlaLeuAlaLeuArgMetTyrAsnGlnGlnInLeuLysArgProArgLeuSerPro 459
DB 2304 GAAGCGGCACCTTGCACAGAAATGATTAACCAACAGATATAAAGCTCCAGTTCTACCC 2363
QY 460 MetGluMetProProValThrSerSerSerTyrSerProIleTyrArgAspArgSerPhe 479
DB 2364 ATGGAATGCCACCAAGTAATCTTCTCATCGTATCTCTCATCTACCGTGATAGAAGCTTT 2423
QY 480 ProSerGlnArgAspAspGlnAspGluIleSerAlaLeuValSerSerTyrLeuGly 499
DB 2424 CCTAGTCAAGAGAGAGATGACCAAGATGAATATCATAGCTCTTGAGTAGTACTCGGC 2483
QY 500 ProSerThrSerPheProHisArgSerArgSerProGluTyrMetValProLeuPro 519
DB 2484 CGGTCAACATCTTTCTCATCGCTCAAGAAGATCCCGGAATATATGGTCCACTTCCA 2543
QY 520 HisGlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSer 539
DB 2544 CATGTGGGTAGGAAGAGTGTATGATGATGAACATCTGGCCCAATTCATATCTCT 2603
QY 540 ProGlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArg 559
DB 2604 CCAGTCCAGGACATAGACTTCATCGACAGTACTCTCGTCTTTGGTTCAAGGACAGAGA 2663
QY 560 HisProLeuGlnTyrSerProProIleHisGlyGlnGlnGlnLeuProTyrGlyIleGln 579
DB 2664 CATCCACTACAGTACTCTCTCCCAATTCATGACAAACAGATTACCATATGGTATACAA 2723
QY 580 ArgValTyrArgHisSerProSerGluGluArgTyrLeuGlyLeuSerAsnGlnArgSer 599
DB 2724 AGGTTTACAGCATCTCCATCTGAAGAAGATATTGGTTTATCCAAATCAAGGTCT 2783
QY 600 ProArgSerAsnSerSerLeuAspProLys 609
DB 2784 CCTGCAAGTAACTCATCTATTAGACCCCAAA 2813

RESULT 7
LOCUS AY198401 3739 bp DNA linear PLN 12-JUN-2003
DEFINITION Arabidopsis thaliana ecotype Shakhara FRIGIDA (FRI) gene, complete cds.
ACCESSION AY198401
VERSION AY198401.1 GI:31558910
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses.
REFERENCE 1 (bases 1 to 3739)
AUTHORS Gazzani,S., Gendall,A.R., Lister,C. and Dean,C.
TITLE Analysis of the molecular basis of flowering time variation in Arabidopsis accessions
JOURNAL Plant Physiol. 132(2), 1107-1114 (2003)
REFERENCE 2 (bases 1 to 3739)
AUTHORS Lister,C. and Dean,C.
TITLE Direct Submission
JOURNAL Submitted (17-DEC-2002) Cell and Developmental Biology, John Innes Center, Colney Lane, Norwich NR4 7UH, UK
FEATURES
source
1..3739
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="4"
/maps="between mi51 and mi122"
/scotype="Shakhara"
<574..>2885
/gene="FRI"
Join(<574..1528,1922..2094,2184..>2885)
/gene="FRI"
/product="FRIGIDA"
Join(574..1528,1922..2094,2184..2885)
CDS
```

```
/gene="FRI"
/note="flowering gene"
/codon_start=1
/product="FRIGIDA"
/protein_id="AAP49807.1"
/db_xref="GI:31558911"
/translation="MSNYPTVAQAQTTPANPLLQHQSEQRRELPRKIVFTSTMSD
ITIGSKQPLKSIDELAASVAETKRFDDIQKHIESIENADSKLSNGVLLA
ANNPHOPMLSPRNNTVETVTVSQSBEIVETSNKPEGRMCLMCKGLRKY
IVANTSDAKLMEETPSALKLAKPAKVLDCIGKFLQGRRAFTKESPMASARQSL
LILEFLMPDRGKGVKLEWIKDEAETAAVAKRLMTGGGLAAAEKMDARGLLL
VACFVPSNFRSTDLLDIRSGNEIAGALKRSQFLVPMVSGIVESIKQMHIEAL
ENVYTFMEDKPSAALVLTSLKMSKESFERAKRQAQSPFLAKFEATKQLAVLSVMQ
CMETKLLDPKMPQWQKEQISLEKDTLQDKEMEEKARSLMESEAALAKRMNQ
QIKRKLSPMPMPVTSYSPIYDRSPFSDDDQDEI SALVSYLGPTSPFHR
RRSPYMWPLPHGGLGRSVYAEHLAPNSYSPGHGRHLRQYSPSLVHGQRHPLYSP
PIHGQQLPYGIQVRVHRSPSEERYVLGLSNQSRPSNSSLDPK"
ORIGIN
Alignment Scores:
Pred. No.: 2,22e-186 Length: 3739
Score: 2975.00 Matches: 607
Percent Similarity: 78.83% Conservative: 0
Best Local Similarity: 78.83% Mismatches: 2
Query Match: 95.84% Indels: 162
DB: 8 Gaps: 2
US-09-890-475-1 (1-609) x AY198401 (1-3739)
QY 1 MetSerAsnTyrProProThrValAlaAlaGlnProThrThrAlaAsnProLeuLeu 20
DB 574 ATGTCCATTAATCCACCGAGCGTGGCGCGCAACCCACACGACGCGAATCCACTGCTG 633
QY 21 GlnArgHisGlnSerGluGlnArgArgGluLeuProLysIleValGluThrGluSer 40
DB 634 CAGCGACATCAATCTGAACGCGACGAGAGAAATACCGAAGATTGTGAAACAGAGTCT 693
QY 41 ThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLysSerIleAsp 60
DB 694 ACAAGTATGGACATTAAGTACGTCATCTAAGCAGCCTCAATTTTGAATCCATAGAC 753
QY 61 GluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAspLeuGln 80
DB 754 GAATTAGTTCGCTTTTTCAGTTGCGAGTGGAAACATTCACACGCGCAATTCATGATCTTCAG 813
QY 81 LysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsnGlyValVal 100
DB 814 AAGCATCATGATCAATCGAAGAACGCAATGATTCGAACTCGAGAGTAAACGCGCTTGTTC 873
QY 101 LeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerProProArgAsnAsnVal 120
DB 874 CTCGCGCGCGGAAACAATAATTTCCATCAGCCGATGTTATCGCTCCGCGGAACAATGTA 933
QY 121 SerValGluThrThrValThrValSerGlnProSerGlnGluIleValProGluThrSer 140
DB 934 TCTGTAGAAACACCGCTCATCTGTGACCCACCGTCTCAGAGATTTGATCCGAGAGAGTCTG 993
QY 141 AsnLysProGluGlyGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyr 160
DB 994 AATAAACCGGAGGCGAGCGATGTTGTGAGTTGATGTTAGCAAGGTTCTGCGTAATATAC 1053
QY 161 IleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAlaLeuLys 180
DB 1054 ATATACCGCAATATCTCTCATCAAGCTAAGTTAATGGAAGAGATTCCTTCAGCTTTGAA 1113
QY 181 LeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGly 200
DB 1114 TTGCGCAAGGAGCCAGCAAGTTTGTATTGGATTGTTTGGCAAGTTTCTTACATCAAGG 1173
QY 201 ArgArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuLys 220
DB 1174 CGTAGAGCATTTACTAAAGAGTCCCTATGAGTCTGCGAGACAAAGTTTTCCTTCTTATA 1233
```

Qy	221	LeuGluSerPheLeuLeuMetProAspArgGlyLysGlyLysValLysIleGluSerTyr	240
Db	1234	CTGAGTCTTTTCTTAATGCTGTATCGTGGTAAGGGAAGGTGAAGATTGAGAGTTGG	1293
Qy	241	IleLysAspGluAlaGluThrAlaAlaValAlaTyrArgLysArgLeuMetThrGluGly	260
Db	1294	ATTAAAGATGAGCGGAGACGGCTGCTGTGCTTGGAGGAAAAGGTTGATGACTGAAGGA	1353
Qy	261	GlyLeuAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPhe	280
Db	1354	GGATTAGCTCGGCTGAGAAAATGGATGTCAGAGGCTTGTCTTTTACGTAGTTGCTGT	1413
Qy	281	GlyValProSerIlePheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySer	300
Db	1414	GGTGTCTCTCAAACTTTAGAGTACAGATTGTCTGGATTGTATGAGATGAGTGTGTTCG	1473
Qy	301	AsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSer	318
Db	1474	AATGAGATTGCGGTGCTTTGAAGCGGTACAGTTTCTGTGCCCTATGGTCTCAGGTACC	1533
Qy	318	-----	318
Db	1534	ATATTGTTCTCACTCGTGAAATTTCATGCCAAAGGTGGTTCCTTTGTTGACATCATC	1593
Qy	318	-----	318
Db	1594	GACCAACATCAAGTTCATCTTCTTTTTTCGATAAGCTTGATGTGTATAACTAGGAGAGC	1653
Qy	318	-----	318
Db	1654	ACATCAAATAATTATAGATGGCAATGACTGATGAGCCAAATCTAGCTAGAAAATTAATCTG	1713
Qy	318	-----	318
Db	1714	GAAAGAACTTGGAACTCTCAACCATAGGTTTTGGTACGAAATTCCTGCTGTCAGAACCA	1773
Qy	318	-----	318
Db	1774	AATGATAGCTATTGCGCTTGAATAGTGTTCCTTTGTTGTTCCAAATATTGGAAGTTAAAA	1833
Qy	318	-----	318
Db	1834	TCATATGACTTAGCTGTGGTATCAATAATTAACTTAAGCTTAAGCAATGCCAAGCTTAAGAAGTGG	1893
Qy	319	-----GlyIleValGluSerSerIleLysArgGlyMet	329
Db	1894	TACTTACAAATATCTATTGCTCATAGGTATGTGAATCAAGTATCAAGCGTGGAAATG	1953
Qy	330	HisIleGluAlaLeuGluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAla	349
Db	1954	CATATTGAAGCTCTTGAGATGGTTATACCTTTGGCATGGAGGATAAGTTTTTCAGTGCT	2013
Qy	350	LeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLys	369
Db	2014	CTAGTTCTTACTTCACTTAAAGATGAGCAAGGAGTCATTTGAGAGGGCAAAACGGAAA	2073
Qy	370	AlaGlnSerProLeuAlaPheLys-----	377
Db	2074	GCCCAGTCACCGCTGGCATTTGT-ATGAACCTTTCCCTTGCACTATTATGTACCTTTTGA	2132
Qy	378	-----GluAla	379
Db	2133	ACTCTTTATCATCATCTGAGTCTGACCATTGATATATTATTTTCTCAACAGAAAGAGCG	2192
Qy	380	AlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeu	399
Db	2193	GCTCAAAAGCAGCTAGCTGTGTTATCATCAGTTATGCAAGTGTATGGAGACTCACAAAGTA	2252
Qy	400	AspProAlaLysGluLeuProGlyTyrGlnIleLysGluGlnIleValSerLeuGluLys	419
Db	2253	GATCTCTCGAAGAACTACCCAGGATGGCAGATCAAGAGCAAAATTTGTAGCTTGGAGAA	2312
Qy	420	AspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGlu	439

Db	2313	GACACTCTTCAGCTCGACAAGAAGAGATGGGAAGAAAGCAAGATCTCTCAGTTTAATGGAG	2372
Qy	440	GlualaAlaLeuAlaLysArgMetTyrAenGlnInlelYArGProArgIeuSerPro	459
Db	2373	GAAGCCGCATTGCCAAGAGAATGTATAACCAACAGATAAACCGTCCAAGGTGTGCACC	2432
Qy	460	MetGluMetProProValThrSerSerSerTyrSerProileTyrArGaAspArgSerPhe	479
Db	2433	ATGGAAATGCCACAGTAACCTCTTCATCGTATTCCTCATCTACCTGATAGAGACTTT	2492
Qy	480	ProSerGlnArgAspAspGlnAspGluileSerAlaLeuValSerSertYrLeuGly	499
Db	2493	CCTAGTCAAAGACGATGACCAAGATGAATATACGTCTCTGTGAGTAGTTACCTCGGC	2552
Qy	500	ProSerThrSerPheProHisArgSerArgArgSerProGluTyrMetValProIeuPro	519
Db	2553	CCGTCAACATCTTTTCTCATCGCTCAAGAAGATCCCCGGAATATATGGTTCCACTTCCA	2612
Qy	520	HisGlycylLeuGlyArgSerValTyrAlaTyrGluHisiLeuAlaProAsnSertYrSer	539
Db	2613	CATGGTGGGTAGGAAGAAGTGATATGTCATATGAACATCTGSCCCCAAAATTCATCTCT	2672
Qy	540	ProGlyHisGlyHisargLeuHisArgGlnTyrSerProSerIeuValHisGlyGlnArg	559
Db	2673	CCAGGTCAACGACATAGACTTCATCGACAGTACTCTCCGTCTTGGTTCAAGCACAGAGA	2732
Qy	560	HieProIeuGlnTyrSerProIleHisGlyGlnGlnLeuProTyrGlyVilLeGln	579
Db	2733	CATCCACTACAGTACTCTCTCCANTTCATGGACAACAACAGTTACCATATGGTATACAA	2792
Qy	580	ArgValTyrArgHisSerProSerGluGluArTyrLeuGlyLeuSerLenGlnArgSer	599
Db	2793	AGGGTTTACAGACATTCACCATCTGAAGAAGAGATATTGGGTTTATTCCAATCAAGGTCT	2852
Qy	600	ProArgSerAsnSerSerIeuAspProLys	609
Db	2853	CCTCGCAGTAACCTCATATTAGACCCCAAA	2882
RESULT 8			
AV198405			
LOCUS			
DEFINITION	Arabidopsis thaliana ecotype Wai-2 FRIGIDA (FRI) gene, complete cds.		
ACCESSION	AY198405		
VERSION	AY198405.1	GI:31558917	
KEYWORDS			
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosoid II; Brassicales; Brassicaceae; Arabidopsia.		
AUTHORS	1 (bases 1 to 3301)		
TITLE	Gazzani,S., Gendall,A.R., Lister,C. and Dean.C. Analysis of the molecular basis of flowering time variation in Arabidopsis accessions		
JOURNAL	Plant Physiol. 132 (2), 1107-1114 (2003)		
REFERENCE	2 (bases 1 to 3301)		
AUTHORS	Gazzani,S. and Dean,C.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-DEC-2002) Cell and Developmental Biology, John Innes Center, Colney Lane, Norwich NR4 7UH, UK		
FEATURES	Location/Qualifiers		
source	1..3301 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /db_xref="taxon:3702" /chromosome="4" /map="between mi51 and mi122" /ecotype="Wai-2" <527..>2838 /gene="FRI"		
gene	join(4527..1481,1875..2047,2137..>2838)		
mRNA			

Db 2206 GATCTCGCAAGAACTACCGAGATGGCAGATCAAGAGCAAAATGTTAGCTGGAGAAA 2265

Qy 420 AspThrLeuGlnLeuAspLysGluMetGluLysAlaArgSerLeuSerLeuMetGlu 439

Db 2266 GACACTCTTCAGCTCGCAAGAGATGGAAGAGAAAGCAAGATCTCTCAGTTTAATGGAG 2325

Qy 440 GluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnLysArgProArgLeuSerPro 459

Db 2326 GAAAGCCGACATGGCAAGAGATGTATACCAACAGATAAACGTCGAAGTTGTCAACC 2385

Qy 460 MetGluMetProProValThrSerSerTyrSerProIleTyrArgAspArgSerPhe 479

Db 2386 ATGGAAATGCCACCACTAATCTTCATCGTATCTCTCTATCTACCGTGATAGAAGCTTT 2445

Qy 480 ProSerGlnArgAspAspGlnAspGlnLysSerAlaLeuValSerSerTyrLeuGly 499

Db 2446 CCTAGTCAAGAGACGATGACCAAGATGAATATATCAGCTCTTGTGAGTAGTACCTCGC 2505

Qy 500 ProSerThrSerPheProHisArgSerArgSerProGluTyrMetValProLeuPro 519

Db 2506 CCGTCAACATCTTTCTCATCGCTCAGAGAGATCCCGGATATATGTTCCACTTCCA 2565

Qy 520 HisGlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSer 539

Db 2566 CATGTGGGTAGGAAGAAGTGTATATGCATATGAACATCTGGGCCCAAAATTCATACTCT 2625

Qy 540 ProGlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArg 559

Db 2626 CCAGTACGAGACATAGACTTCATCGACAGACTCTCGTCTTTGGTTCCAGCAAGAGA 2685

Qy 560 HisProLeuGlnTyrSerProProIleHisGlyGlnGlnLeuProTyrGlyLeuGln 579

Db 2686 CATCCACTACAGTACTCTCTCCATCCATTCATGACAAACAGTATACCATATGGTATACAA 2745

Qy 580 ArgValTyrArgHisSerProSerGluGluArgTyrLeuGlyLeuSerAsnGlnArgSer 599

Db 2746 AGGGTTACAGACATTCACCATCTGAGAAAGATATTTGGGTATTCATCAATCAAGGTCT 2805

Qy 600 ProArgSerAsnSerSerLeuAspProlys 609

Db 2806 CCTCGCAGTAATCATCATTAGACCCCAAA 2835

RESULT 9

AY198404

LOCUS

DEFINITION Arabidopsis thaliana ecotype Cvi FRIGIDA-like protein gene, complete sequence.

ACCESSION AY198404

VERSION AY198404.1 GI:31558916

KEYWORDS

SOURCE

ORGANISM Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 3758)

Gazzani, S., Gendall, A.R., Lister, C. and Dean, C.

Analysis of the molecular basis of flowering time variation in Arabidopsis accessions

Plant Physiol. 132 (2), 1107-1114 (2003)

2 (bases 1 to 3758)

Gazzani, S. and Dean, C.

Direct Submission

Submitted (17-DEC-2002) Cell and Developmental Biology, John Innes Center, Colney Lane, Norwich NR4 7UH, UK

Location/Qualifiers

1. 3758

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/db_xref="taxon:3702"

/chromosome="4"

/map="between mi51 and mi122"

/ecotype="Cvi"

FEATURES

source

misc_feature 577..2888

/note="similar to FRIGIDA (FRI)"

ORIGIN

Alignment Scores: 9.92e-185 Length: 3758

Prod. No.: 2950.00 Matches: 604

Score: 78.57% Conservattive: 1

Best Local Similarity: 78.44% Mismatches: 4

Query Match: 95.04% Indels: 162

DB: 8 Gaps: 2

US-09-890-475-1 (1-609) x AY198404 (1-3758)

Qy 1 MetSerAsnTyrProProThrValAlaAlaGlnProThrThrAlaAsnProLeuLeu 20

Db 577 ATGTCCAAATATTCACCGACGCTGGCGGCGCAAAACCAACACGCGCGAATCCACTGCTG 636

Qy 21 GlnArgHisGlnSerGluGlnArgArgArgGluLeuProLysIleValGluThrGluSer 40

Db 637 CAGCGACATCAATCTGAACAGCGACGAGAGAAATACCGAAGATTGTGCAAAACAGAGTCT 696

Qy 41 ThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLysSerIleAsp 60

Db 697 ACAAGTATGAGCATTCAGCATCGTCAATCTAAGCAGCGCTCAATTTTGAATCCATAGAC 756

Qy 61 GluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAspAspLeuGln 80

Db 757 GAATTAGTCGCTTTTCAGTTCAGTGGAAACATTCMAATGCCAATTCGATGATCTTCAG 816

Qy 81 LysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsnGlyValVal 100

Db 817 AAGCATCATCGAGTCAATCGAAACGCAATTTGATTCAAACTCGAGAGTAACCGCGTTGC 876

Qy 101 LeuAlaAlaArgAsnAsnPheHisGlnProMetLeuSerProProArgAsnAsnVal 120

Db 877 CTGCGCGCGGGAACAAATTAATTTCCATCAGCCGATGTTATCGCTCCGCGGAAACAATGTA 936

Qy 121 SerValGluThrThrValThrValSerGlnProSerGlnGluIleValProGluThrSer 140

Db 937 TCTGTAGAAACCAACCGCTCTGTAGCCCAACCGTCTCAGAGATTGTACCGAGAGCGTCG 996

Qy 141 AsnLysProGluGlyGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyr 160

Db 997 AATAAACCGGAGCGGAGCGATGTGTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1056

Qy 161 IleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAlaLeuLys 180

Db 1057 ATATACGCGAATATCTCTGCAACAGCTAAGTTAATGGAAGAGATTCTCTTCAGCTTTGAAA 1116

Qy 181 LeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGly 200

Db 1117 TTGGCCCAAGGAGCCAGCGAAGTTTGTATTGGATTGTATTGGCAAGTTTACTTACAAGGG 1176

Qy 201 ArgArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuIle 220

Db 1177 CGTAGAGCAATTTACTAAAGATCGCTATGAGCTCTGCGAGACAGATTTCGCTTCTTATA 1236

Qy 221 LeuGluSerPheLeuLeuMetProAspArgGlyLysGlyLysValLysIleGluSerTyr 240

Db 1237 CTGGAGTCTTTTCTTCTAATGCTGATCGTGTGTTAAGGGAAGGTGAAGATTGAGAGTGTG 1296

Qy 241 IleLysAspGluAlaGluThrAlaAlaValAlaAlaTyrArgLysArgLeuMetThrGluGly 260

Db 1297 AITPAAGATGAGCGGAGACGCTGCTGTGTTGTTGGAGGAAGAGTTGATGACTGAGGA 1356

Qy 261 GlyLeuAlaAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPhe 280

Db 1357 GGATTAGTCGCGCTCAGAAATGGATGCAAGGGGTTTCTTTACTAGTTGCTTCTTTT 1416

Qy 281 GlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySer 300

Db 1417 GGTGTTTCTTCAAACTTTAGGAGTACAGATTTGCTGGATTTGATAGGAGTAGTGTTGTCG 1476

QY	301	AspGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSer	318
DB	1477	AATGAGATTGCCGGTGCTTTTGAAGCGGTACACAGTCTTCTTGCTCCCTATGGTCTCAGGTACC	1536
QY	318	-----	318
DB	1537	ATATTCGTTCTCACTCGGTGAATTCATTGCAAGGTGGTTCCTTTGTGTGACATCATC	1596
QY	318	-----	318
DB	1597	GACCAACATCAAGTTCCATCTTTGTGTTTTTCGATAAGCTTGATGGTATAAACTAGGAGAGC	1656
QY	318	-----	318
DB	1657	ACATCAAAATATTAGAGTGCAAATGACTGATTGAGCCAAATCCTAGCTAGAAATTAATCTG	1716
QY	318	-----	318
DB	1717	GAAGAACCTTGGAACTCTCAACCATAGGTTTTGGTACGAAATTTGCTTGTCAGAACCA	1776
QY	318	-----	318
DB	1777	AATGATAGGCTATTGCGCTTGAAATAGTGTTCCTTGCTGGTTTCCAAATATTGGAAAGTTAAAA	1836
QY	318	-----	318
DB	1837	TCATATGACTTAGCTGTGGTACTAATAATTAAGCTTAAGCAATGCCAACTCTAAGAAGTGG	1896
QY	319	-----GlyIleValGluSerSerIleLysArgGlyMet	329
DB	1897	TACTTACACAATATTCTATTGGTCATAGGTATAGTTGAATCAAGTATCAAGCTGGGAATG	1956
QY	330	HisIleGluAlaLeuGluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAla	349
DB	1957	CATATTGAAGCTCTTGAGATGGTTTATACCTTTGGCATCGAGGATAAGTTTTCAGCTGCT	2016
QY	350	LeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLys	369
DB	2017	CTAGTTCTAACTTCATTCTTAAAGATGAGCAAGAGGTCATTTGAGAGGGCAAAACGGAAA	2076
QY	370	AlaGlnSerProLeuAlaPheLys-----	377
DB	2077	GCCAGTCACGGCTGGCATTTGT-ATGAACCCCTCCCTTGACATATTGACCTTTATGA	2135
QY	378	-----GluAla	379
DB	2136	ACTCTTTATCATCTCTGAGTCTGACCATTGATATATTATTTCTCAACAGCAAGAGCG	2195
QY	380	AlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeu	399
DB	2196	GCTACAAGCAGCTAGCTGTGTATCATCATGTTATGCAGTGTATGGAGACTCACAGTTA	2255
QY	400	AspProAlaLysGluLeuProGlyTTPGlnLleLysGluInleValSerLeuGluLys	419
DB	2256	GATCCTCGAAGAAGCACTACCAAGATGCGAGATGAGAGAGCAAAATTTGTTAGCTTTGGAGAAA	2315
QY	420	AspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGlu	439
DB	2316	GACACTCTTCAGCTCGAACAAGAGATGGAGAGAGCAAGATCTCTCAGTTTAATGGAG	2375
QY	440	GluAlaAlaLeuAlaLysArgMetLysAsnGlnGlnLleLysArgProArgLeuSerPro	459
DB	2376	GAAGCCGCATTTGCGAAGAGAAATGTATAACCAACAGATAAAACGTCACAGGTTGTCAACC	2435
QY	460	MetGluMetProProValThrSerSerSerTyrSerProIleTyrArgAspArgSerPhe	479
DB	2436	ATGGAAATGCCACAGTAACCTTCTTCATCGTATTCTCTATCTACCGTGATAGAGCTTT	2495
QY	480	ProSerGlnArgAspAspGlnAspGluIleSerAlaLeuValSerSerTyrLeuGly	499
DB	2496	CCTAGTCAAGAGACGATGACCAAGATGAATATCAGCTCTTGTGAGTAGTACTTCGCGC	2555

Qy	500	ProSerThrSerPheProHisArgSerArgArgSerProGluTyrMetValProLeuPro	519
Db	2556	CCGTCAACATCTTTTCTCTCATCGCTCARGAAGATCCCCGAATATATGGTTCACATTCCA	2615
Qy	520	HisGlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSer	539
Db	2616	CATGGTGGGTAGGAAGAAGTGATATGCATATGAACATCTGGCCCCAAATTTCATACTCT	2675
Qy	540	ProGlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArg	559
Db	2676	CCAGGTCACGGACATAGACTTCATGCACAGTACTCTCCGTCCTTGGTTTCAGGACACAGA	2735
Qy	560	HisProLeuGlnTyrSerProProlIleHisGlyGlnGlnLeuProTyrGlyIleGln	579
Db	2736	CATCCACTACAGTACTCTCTCCCAATTCATGGACACACACAGTACCATATGTTATACAA	2795
Qy	580	ArgValTyrArgHisSerProSerGluGluuArgTyrLeuGlyLeuSerArgGlnArgSer	599
Db	2796	AGGGTTACAGACATTCACCATCTGAAGAAGAAGATATTTGGGTTTATCCAAATCAAGGTCT	2855
Qy	600	ProArgSerAsnSerSerLeuAspProLys	609
Db	2856	CTCTCGCAGTAACATCATATTAGACCCCAA	2885

RESULT 10

Accession	LOCUS	Size	Library	Strain	Accession
F6N23	LOCUS	91849 bp	DNA	linear	PLN 12-NOV-1999
F6N23	DEFINITION	Arabidopsis thaliana BAC F6N23.			

AF036919.2 GI:0362042
 .
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
REFERENCE
AUTHORS
TITLE
1 (bases 1 to 91849)
Geisel, C.
The sequence of A. thaliana F6N23
Unpublished
2 (bases 1 to 91849)
Washington University Genome Sequencing Center.
The A. thaliana Genome Sequencing Project
Unpublished

MAPPING: Clones were assigned to the YAC map by hybridization by M. Lodhi, Cold Spring Harbor Laboratories, and fingerprinted by M. Marra, WashU, to pick the best candidates for sequencing.

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted:

all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone

NEIGHBORING COSMID INFORMATION:

The 5' clone is F5110, 200 bp overlap; 3' clone is F15P23, 900 bp overlap. Actual start of this clone is at base position 104960 of CELF5110; actual end is at 91149 of CELF6N23

The clone sequenced to the left is F5110. The actual start of this clone is at base position 104960 of F5110; actual end is at base position 91149 of F6N23.

NOTES:

Coding sequences below are predicted from computer analysis, using the program Genefinder (P. Green and L. Hillier, ms in preparation). Location/Qualifiers

FEATURES

```

source
1. .91849
  /organism="Arabidopsis thaliana"
  /mol_type="genomic DNA"
  /cultivar="Columbia"
  /db_xref="taxon:3702"
  /chromosome="IV"
  /map="unknown"
  /clones="F6N23"
  /complement(187. .907)
  /gene="F6N23.23"
  /complement(join(187. .347,396. .512,598. .907))
  /gene="F6N23.23"
  /codon_start=1
  /evidence=not_experimental
  /protein_id="AAC13624.2"
  /db_xref="GI:6382044"

gene
CDS
  /translation="MENLSPDENDHOSPRSDONDSOAVETPVYSTMSIDSFVYPR
  TCSFTSGSDIDETMSRCSEASPCNWPVLTSKSKKCLSGLEMSQNECLVQGEISE
  PELTMRERAKLLGDMGSGKGVCTATISNATNLNDKNDTSFNDVCSYVW
  TESVRAVRNEESIVBERNELSFICMLHCRHP"
  /complement(4919. .7497)
  /gene="F6N23.22"
  /complement(join(4919. .4993,5078. .5488,5567. .6055,
  6157. .6213,6498. .6594,6935. .7222,7349. .7497))
  /gene="F6N23.22"
  /codon_start=1
  /evidence=not_experimental
  /protein_id="AAC13623.1"
  /db_xref="GI:3047112"
  /translation="MSLTMDGVAAAGRSKQNSLLRKLALAVRSVQMSYAFWS
  SLTQGVLEMGEGYGMKKRKYSHYKGLQSKELRKLKYLMLGDSGTWTST
  THNLNDDNCHSTSMWLDLSDEWYVLSMSVFPSPCLPGASATGETIWL
  CNAOYAEKILRSLLARTVCPYLAGVIELGYTELISRDHNLRLNKSCLMEISAH
  QNDEKKNKIKISEEKQLPLGLSDEDLHYKRLISVLYNSADRSGKNDKIRHRCP
  NIVTSEPSSFLRWQCEQQVGFQKKSQNVLRKILHDLVPLMKRMFPQSGNLN
  QDDPSDRKNEKESVLTWPTVNEVDKESILNNTIKYQLEEARVEELSCWSVN
  FVQRKTEINLNDLSITSGNYDDSKIDDSGETQVTPFRDKTHLRKIKETE
  VVIEKRCYROYIVADIMETLSNLHMDAFSVRSHTLNKFLTLNLKAKFRGAASVGM
  IKRLRRVIGDLF"
  /complement(12675. .12675)
  /gene="F6N23.1"
  /complement(join(10235. .10495,10693. .10896,10998. .11192,11341. .11505,
  11633. .11893,11998. .12216,12315. .12534,12617. .12675))
  /gene="F6N23.1"
  /note="similar to the family of glycosyl hydrolases"
  /codon_start=1
  /evidence=not_experimental
  /protein_id="AAC13634.1"
  /db_xref="GI:3047123"
  /translation="MAIRLNHSVIPSVYKLGAPTRVSGARSLPFSVGDWGRGVSTFGA
  RPLVLAKVLAESTEDRVPIDDDDSDTQLSVSEFHLDDSLILGVIDMNSSEVVEPEE
```

gene

```

CDS
  /complement(join(13027. .13335,13480. .13615,13712. .13840,
  13943. .14001,14087. .14791))
  /gene="F6N23.21"
  /codon_start=1
  /evidence=not_experimental
  /protein_id="AAC13622.1"
  /db_xref="GI:3047111"
  /translation="MSILCCVPVLECYVLCGTWLMKKCLYSAGHSESNWGLATSDE
  FEPFRIICRLIAYVEENLHDPWAPDGGIDENHVLKDYDQTEGRVTPMYILD
  HENGDVLAIRGLNAKECDYAVLLDNKLGQTFDGGVYVHNLKAAWVPEEBHVVL
  RELLEANPSYSLTFVGHSLGAGVSLVLFVIONRVLGNIERKRIRCFATAPRCMS
  LHLAVYADVNSVLQDDPLPTTALENVFKSIIICLLCLCTCKDTPFERKL
  KDARLYAPRLYHIVKRLGRYPPVATVPDGRFEQIVLSNATADHAIWI
  ERESORALDKIVRQKSIQVEDHDEYRAIMKAASINLPMSPPSIGTFHDTREGESEA
  GSGMEGSPSGWSFKGMRKRKQDFIDCHFPVNDNSEHMIIFKNQESQALLCGOET"
  /complement(15539. .16687)
  /gene="F6N23.2"
  /complement(join(15539. .15554,15672. .15764,15953. .16026,16133. .16208,
  16307. .16353,16466. .16687))
  /gene="F6N23.2"
  /note="Similar to protein kinase; coded for by A. thaliana
  cDNA H36947; coded for by A. thaliana cDNA H37158"
  /codon_start=1
  /protein_id="AAC13612.1"
  /db_xref="GI:3047101"
  /translation="MNTESVVEFLGNVTLORLPSSSLKRISEVVVKYGDYVVR
  ENQVDGVIFLLQGVRAQVLAAGEENYQEPFLKRYDFGFG:FGDYVSADYVATTE
  LTCLLMSDHRALLKISVSDSKERCLVDELLYDELDVRSVSPNCPSPSSISVLMIK
  VLTIFVYHHSIESEL"
  /complement(16976. .19151)
  /gene="F6N23.3"
  /complement(join(16976. .17111,17262. .17415,17785. .17889,18316. .18417,
  18618. .18666,19056. .19151))
  /gene="F6N23.3"
  /codon_start=1
  /evidence=not_experimental
  /protein_id="AAC13635.1"
  /db_xref="GI:3047124"
  /translation="MEIYSNHCDTTVYASQALAAASKVTVMKI VNHFCYFLLVGDI
  NPIIYDYNLRDGNPATRSVDAPKGTITLTFASQVSCYIPEIHFESLDVRNQ
  VATEISVPPIDIRFCEPNRSYKQNSKPRCVAVASDLIFATISLPHRSEGMVA
  LSLSHMPHPRVPRADMLLVIVISPTATESGFAIGKFNKNGEVRCIEBE"
  /complement(20843. .21366)
  /gene="F6N23.20"
  /complement(join(20843. .20944,21002. .21056,21103. .21155,
  21292. .21366))
  /gene="F6N23.20"
  /codon_start=1
  /evidence=not_experimental
  /protein_id="AAC13621.1"
  /db_xref="GI:3047110"
  /translation="MVATAVASYFLITGDYGPENVLDPYLSMILLIVEYLIRMF
  SYWVMLHLFVLIHVIHWIKRLAAQDSVKQIFGPKGSPSGKESDTAK"
  /complement(22410. .24753)
  /gene="F6N23.19"
  /complement(join(22410. .22521,23220. .23538,23620. .23939,
  24009. .24137,24240. .24352,24439. .24548,24639. .24753))
  /gene="F6N23.19"
  /codon_start=1
  /evidence=not_experimental
  /protein_id="AAC13637.1"
  /db_xref="GI:3047126"
  /translation="MSCPSYINVSUKPFFHLMQDQLITNVVKYQGRNWKRTAECULFG
```

gene

```

CDS
  /complement(join(15539. .15554,15672. .15764,15953. .16026,16133. .16208,
  16307. .16353,16466. .16687))
  /gene="F6N23.2"
  /note="Similar to protein kinase; coded for by A. thaliana
  cDNA H36947; coded for by A. thaliana cDNA H37158"
  /codon_start=1
  /protein_id="AAC13612.1"
  /db_xref="GI:3047101"
  /translation="MNTESVVEFLGNVTLORLPSSSLKRISEVVVKYGDYVVR
  ENQVDGVIFLLQGVRAQVLAAGEENYQEPFLKRYDFGFG:FGDYVSADYVATTE
  LTCLLMSDHRALLKISVSDSKERCLVDELLYDELDVRSVSPNCPSPSSISVLMIK
  VLTIFVYHHSIESEL"
  /complement(16976. .19151)
  /gene="F6N23.3"
  /complement(join(16976. .17111,17262. .17415,17785. .17889,18316. .18417,
  18618. .18666,19056. .19151))
  /gene="F6N23.3"
  /codon_start=1
  /evidence=not_experimental
  /protein_id="AAC13635.1"
  /db_xref="GI:3047124"
  /translation="MEIYSNHCDTTVYASQALAAASKVTVMKI VNHFCYFLLVGDI
  NPIIYDYNLRDGNPATRSVDAPKGTITLTFASQVSCYIPEIHFESLDVRNQ
  VATEISVPPIDIRFCEPNRSYKQNSKPRCVAVASDLIFATISLPHRSEGMVA
  LSLSHMPHPRVPRADMLLVIVISPTATESGFAIGKFNKNGEVRCIEBE"
  /complement(20843. .21366)
  /gene="F6N23.20"
  /complement(join(20843. .20944,21002. .21056,21103. .21155,
  21292. .21366))
  /gene="F6N23.20"
  /codon_start=1
  /evidence=not_experimental
  /protein_id="AAC13621.1"
  /db_xref="GI:3047110"
  /translation="MVATAVASYFLITGDYGPENVLDPYLSMILLIVEYLIRMF
  SYWVMLHLFVLIHVIHWIKRLAAQDSVKQIFGPKGSPSGKESDTAK"
  /complement(22410. .24753)
  /gene="F6N23.19"
  /complement(join(22410. .22521,23220. .23538,23620. .23939,
  24009. .24137,24240. .24352,24439. .24548,24639. .24753))
  /gene="F6N23.19"
  /codon_start=1
  /evidence=not_experimental
  /protein_id="AAC13637.1"
  /db_xref="GI:3047126"
  /translation="MSCPSYINVSUKPFFHLMQDQLITNVVKYQGRNWKRTAECULFG
```

gene

```

CDS
  /complement(join(15539. .15554,15672. .15764,15953. .16026,16133. .16208,
  16307. .16353,16466. .16687))
  /gene="F6N23.2"
  /note="Similar to protein kinase; coded for by A. thaliana
  cDNA H36947; coded for by A. thaliana cDNA H37158"
  /codon_start=1
  /protein_id="AAC13612.1"
  /db_xref="GI:3047101"
  /translation="MNTESVVEFLGNVTLORLPSSSLKRISEVVVKYGDYVVR
  ENQVDGVIFLLQGVRAQVLAAGEENYQEPFLKRYDFGFG:FGDYVSADYVATTE
  LTCLLMSDHRALLKISVSDSKERCLVDELLYDELDVRSVSPNCPSPSSISVLMIK
  VLTIFVYHHSIESEL"
  /complement(16976. .19151)
  /gene="F6N23.3"
  /complement(join(16976. .17111,17262. .17415,17785. .17889,18316. .18417,
  18618. .18666,19056. .19151))
  /gene="F6N23.3"
  /codon_start=1
  /evidence=not_experimental
  /protein_id="AAC13635.1"
  /db_xref="GI:3047124"
  /translation="MEIYSNHCDTTVYASQALAAASKVTVMKI VNHFCYFLLVGDI
  NPIIYDYNLRDGNPATRSVDAPKGTITLTFASQVSCYIPEIHFESLDVRNQ
  VATEISVPPIDIRFCEPNRSYKQNSKPRCVAVASDLIFATISLPHRSEGMVA
  LSLSHMPHPRVPRADMLLVIVISPTATESGFAIGKFNKNGEVRCIEBE"
  /complement(20843. .21366)
  /gene="F6N23.20"
  /complement(join(20843. .20944,21002. .21056,21103. .21155,
  21292. .21366))
  /gene="F6N23.20"
  /codon_start=1
  /evidence=not_experimental
  /protein_id="AAC13621.1"
  /db_xref="GI:3047110"
  /translation="MVATAVASYFLITGDYGPENVLDPYLSMILLIVEYLIRMF
  SYWVMLHLFVLIHVIHWIKRLAAQDSVKQIFGPKGSPSGKESDTAK"
  /complement(22410. .24753)
  /gene="F6N23.19"
  /complement(join(22410. .22521,23220. .23538,23620. .23939,
  24009. .24137,24240. .24352,24439. .24548,24639. .24753))
  /gene="F6N23.19"
  /codon_start=1
  /evidence=not_experimental
  /protein_id="AAC13637.1"
  /db_xref="GI:3047126"
  /translation="MSCPSYINVSUKPFFHLMQDQLITNVVKYQGRNWKRTAECULFG
```

gene

```

CDS
  /complement(join(15539. .15554,15672. .15764,15953. .16026,16133. .16208,
  16307. .16353,16466. .16687))
  /gene="F6N23.2"
  /note="Similar to protein kinase; coded for by A. thaliana
  cDNA H36947; coded for by A. thaliana cDNA H37158"
  /codon_start=1
  /protein_id="AAC13612.1"
  /db_xref="GI:3047101"
  /translation="MNTESVVEFLGNVTLORLPSSSLKRISEVVVKYGDYVVR
  ENQVDGVIFLLQGVRAQVLAAGEENYQEPFLKRYDFGFG:FGDYVSADYVATTE
  LTCLLMSDHRALLKISVSDSKERCLVDELLYDELDVRSVSPNCPSPSSISVLMIK
  VLTIFVYHHSIESEL"
  /complement(16976. .19151)
  /gene="F6N23.3"
  /complement(join(16976. .17111,17262. .17415,17785. .17889,18316. .18417,
  18618. .18666,19056. .19151))
  /gene="F6N23.3"
  /codon_start=1
  /evidence=not_experimental
  /protein_id="AAC13635.1"
  /db_xref="GI:3047124"
  /translation="MEIYSNHCDTTVYASQALAAASKVTVMKI VNHFCYFLLVGDI
  NPIIYDYNLRDGNPATRSVDAPKGTITLTFASQVSCYIPEIHFESLDVRNQ
  VATEISVPPIDIRFCEPNRSYKQNSKPRCVAVASDLIFATISLPHRSEGMVA
  LSLSHMPHPRVPRADMLLVIVISPTATESGFAIGKFNKNGEVRCIEBE"
  /complement(20843. .21366)
  /gene="F6N23.20"
  /complement(join(20843. .20944,21002. .21056,21103. .21155,
  21292. .21366))
  /gene="F6N23.20"
  /codon_start=1
  /evidence=not_experimental
  /protein_id="AAC13621.1"
  /db_xref="GI:3047110"
  /translation="MVATAVASYFLITGDYGPENVLDPYLSMILLIVEYLIRMF
  SYWVMLHLFVLIHVIHWIKRLAAQDSVKQIFGPKGSPSGKESDTAK"
  /complement(22410. .24753)
  /gene="F6N23.19"
  /complement(join(22410. .22521,23220. .23538,23620. .23939,
  24009. .24137,24240. .24352,24439. .24548,24639. .24753))
  /gene="F6N23.19"
  /codon_start=1
  /evidence=not_experimental
  /protein_id="AAC13637.1"
  /db_xref="GI:3047126"
  /translation="MSCPSYINVSUKPFFHLMQDQLITNVVKYQGRNWKRTAECULFG
```

gene

```

CDS
  /complement(join(15539. .15554,15672. .15764,15953. .16026,16133. .16208,
  16307. .16353,16466. .16687))
  /gene="F6N23.2"
  /note="Similar to protein kinase; coded for by A. thaliana
  cDNA H36947; coded for by A. thaliana cDNA H37158"
  /codon_start=1
  /protein_id="AAC13612.1"
  /db_xref="GI:3047101"
  /translation="MNTESVVEFLGNVTLORLPSSSLKRISEVVVKYGDYVVR
  ENQVDGVIFLLQGVRAQVLAAGEENYQEPFLKRYDFGFG:FGDYVSADYVATTE
  LTCLLMSDHRALLKISVSDSKERCLVDELLYDELDVRSVSPNCPSPSSISVLMIK
  VLTIFVYHHSIESEL"
  /complement(16976. .19151)
  /gene="F6N23.3"
  /complement(join(16976. .17111,17262. .17415,17785. .17889,18316. .18417,
  18618. .18666,19056. .19151))
  /gene="F6N23.3"
  /codon_start=1
  /evidence=not_experimental
  /protein_id="AAC13635.1"
  /db_xref="GI:3047124"
  /translation="MEIYSNHCDTTVYASQALAAASKVTVMKI VNHFCYFLLVGDI
  NPIIYDYNLRDGNPATRSVDAPKGTITLTFASQVSCYIPEIHFESLDVRNQ
  VATEISVPPIDIRFCEPNRSYKQNSKPRCVAVASDLIFATISLPHRSEGMVA
  LSLSHMPHPRVPRADMLLVIVISPTATESGFAIGKFNKNGEVRCIEBE"
  /complement(20843. .21366)
  /gene="F6N23.20"
  /complement(join(20843. .20944,21002. .21056,21103. .21155,
  21292. .21366))
  /gene="F6N23.20"
  /codon_start=1
  /evidence=not_experimental
  /protein_id="AAC13621.1"
  /db_xref="GI:3047110"
  /translation="MVATAVASYFLITGDYGPENVLDPYLSMILLIVEYLIRMF
  SYWVMLHLFVLIHVIHWIKRLAAQDSVKQIFGPKGSPSGKESDTAK"
  /complement(22410. .24753)
  /gene="F6N23.19"
  /complement(join(22410. .22521,23220. .23538,23620. .23939,
  24009. .24137,24240. .24352,24439. .24548,24639. .24753))
  /gene="F6N23.19"
  /codon_start=1
  /evidence=not_experimental
  /protein_id="AAC13637.1"
  /db_xref="GI:3047126"
  /translation="MSCPSYINVSUKPFFHLMQDQLITNVVKYQGRNWKRTAECULFG
```

SENRNDVQCHRWLKVLDPSLOKQWKEDELLSELVYKDYMNDRPPWKSLSKEL
 PGRICKCREWHNHLNPTIIKSPWREZEELIVQAKQNGKNKWAELAKLDPRTENN
 IKHNWCVKRLQFPNLSFGVYVYKSPSSGFNFQORNTWESCIISQIKEAA
 KSPORDFDLTLGLNWSISSTSLRGSESVSSVDSVCARLNACLETQNSNDIV
 CHVEKREKRLWAARTFDTPSIISKTSPASGLKRLQKDYDTPFTDARSHMSSE
 DVSVAAPSSKVRFRVQRNCTGSKXPLERLDFLLWDEHRRNGIYNFVSRIPLQKS
 DLKSLVLPFWLR"

Alignment Scores:

Pred. No.: 5,496-182 Length: 91849
 Score: 2936.00 Matches: 603
 Percent Similarity: 78.95% Conservative: 1
 Best Local Similarity: 78.82% Mismatches: 5
 Query Match: 94.53% Indels: 158
 DB: 8 Gaps: 2

US-09-890-475-1 (1-609) x F6N23 (1-91849)

Qy	1	MetSerAsnTyrProProThrValAlaAlaGlnProThrThrAlaAsnProLeuLeu	20
Db	56839	ATGTCCAATTATCCACCGAGGTGGCGGCAACCCACACGCGGAAATCCACTGCTG	56898
Qy	21	GlnArgHisGlnSerGluGlnArgArgGluLeuProLysIleValGluThrGluSer	40
Db	56899	CAGCGACATCAATCTGAACAGCGACGAGAGAAATACCGAAGATTGTGGAACAGAGTCT	56958
Qy	41	ThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLysSerIleAsp	60
Db	56959	ACAAGTATGGACATTACGATCGTCAATCTAATGACGACGCTCAATTTTGAATCCATAGAC	57018
Qy	61	GluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAspAspLeuGln	80
Db	57019	GAATTAGTGGCTGCTTTTCAGTTGCACTGGAAACATTCACAAACCCCAATTCGATGATCTTCAG	57078
Qy	81	LysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsnGlyValVal	100
Db	57079	PAGCACATCGACTCAATCGAANACCAATGATTCACAACTCGAGATTAACGGGTTGTC	57138
Qy	101	LeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerProArgAsnVal	120
Db	57139	CTCGCGCGCGGCAACAATAATTTCCATCGCCGATGTTATCGCTCCCGCGGAACAATGTA	57198
Qy	121	SerValGluThrThrValThrValSerGlnProSerGlnGluIleValProGluThrSer	140
Db	57199	TTCTGTAGAAACCCCGTACTGTGAGCCACACCGTCTCAGAGATTGTACCGGAGACGTCG	57258
Qy	141	AsnLysProGluGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyr	160
Db	57259	ATAAACCGAGCGGGAACGTATATGTAGTTGATGTGTAGCAAAAGGCTCGCTAAATAC	57318
Qy	161	IleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAlaLeuLys	180
Db	57319	ATATACGCGAATATCTGATCAAGCTAAGTTAATGGAAGAGATTCCTTCAGCTTTGANA	57378
Qy	181	LeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGly	200
Db	57379	TTGGCAAGAGCCAGCGAAGTTTGTATTGGATTGTATTGGCAAGTTTACTTCAACGGG	57438
Qy	201	ArgArgAlaPheThrLysGlnSerProMetSerSerAlaArgGlnValSerLeuLeuIle	220
Db	57439	CGTAGAGCAATTTACTAAAGAGTCGCTATGAGCTCTGCGAGACAAGTTTCGCTTCTTATA	57498
Qy	221	LeuGluSerPheLeuLeuMetProAspArgGlyLysGlyLysValIleGluSerTyr	240
Db	57499	CTGGAGTCTTTCTTCTAATGCTGATCGTGTGTAAAGGAGAGTTCGAGATTGG	57558
Qy	241	IleLysAspGluAlaGluThrAlaValAlaValIleArgLysArgLeuMetThrGluGly	260
Db	57559	ATTAAGATGAGCGGAGACGGCTGCTGTGCTTGGAGAAAAGTTTGTATGACTGAAGGA	57618
Qy	261	GlyLeuAlaAlaLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPhe	280
Db	57619	GGATTAGCTCGCGCTGAGAAATGATGCAAGGGGTTTGTCTTACTAGTTGCTTGT	57678

Qy	281	GlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySer	300
Db	57679	GGTGTCTTCAAACTTTAGGAGTACAGATTTGCTGGATTTGATAGGATGAGTGCTCG	57738
Qy	301	AsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMet	316
Db	57739	AATGAGATTGCCGGTGTCTTTGAAGCGGTACAGTTTTCAGGTACCATA-TTCTGTCTCTCAC	57797
Qy	316	-----	316
Db	57798	TGGTGAATTCATTCGCAAAAGTGGTTCCTTTTGTGACATCATCGACCAATCAAGTT	57857
Qy	316	-----	316
Db	57858	CCATCTTTGTTTTTCGATAAGCTTGATGATATAAATAGGAGAGCACATCAATATTAG	57917
Qy	316	-----	316
Db	57918	AGTGAATGACTGATTGAGCCAAATCTTAGCTAGAAATTAATCTGGAAGAAGACTTGGAAC	57977
Qy	316	-----	316
Db	57978	TCTCAACCATAGTGTTCGTACGAAATTCGTTGCTGCAGAACCAATGATAGGCTATTG	58037
Qy	316	-----	316
Db	58038	CCTTGAATAGTGTTCCTTTGCTTTCCAAATATGGAAGTTAAATCATATGACTTAGCT	58097
Qy	316	-----	316
Db	58098	GTTGGATCTAATTAAGCTTACCAATGCAATGCAACTTAAGAAGTGGTACTTACACAATAT	58157
Qy	317	-----	317
Db	58158	CTATTGGTTCATAGTATAGTTGAATCAAGTATCAAGCTGGAATGATATGAAAGCTCT	58217
Qy	335	GluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSer	354
Db	58218	GAATGGTTTATACCTTTGGCATGGAGGATAGTTTTCAGTCTCTAGTTCTTACTTCA	58277
Qy	355	PheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeu	374
Db	58278	TTCTTAAAGATGAGCAAGGAGTCATTTGAGAGCGCAAAACGGAAGCCAGTCACCGCTG	58337
Qy	375	Alaphelys-----	377
Db	58338	GCATTTGT-ATGAACCCCTTCCCTTGCACTATGTACTTTTGAAGTCTTTATCATCAT	58396
Qy	378	-----	378
Db	58397	CTGAGCTGACCATTCATATATTTTCTCAACAGAAAGAGCGCTACAAAGCAGCTA	58456
Qy	385	AlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeuAspProAlaLysGlu	404
Db	58457	GCTGTGTATCATCAGTTATGCACTATGAGAGACTCACAAGTTAGATCCTCGCAAGAA	58516
Qy	405	LeuProGlyTyrGlnIleLysGluGlnIleValSerLeuGluLysAspThrLeuGlnLeu	424
Db	58517	CTACCAAGATGCGAGATCAAGAGCAAAATTTGTTAGCTTGGAGAAAGACACTCTTCACTC	58576
Qy	425	AspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGluAlaLeuAla	444
Db	58577	GACAAAGATGGAAGAGAAAGCAAGATCTCTCAGTTTAAATGAGGAAAGCGGCTTGGC	58636
Qy	445	LysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSerProMetGluMetProPro	464
Db	58637	AAGAGATGTATACCAACAGATAAACGCTCCAAAGGTTATCACCCTGGAATGCCACCA	58696
Qy	465	ValThrSerSerSerTyrSerProIleTyrArgAspArgSerPheProSerGlnArgAsp	484
Db	58697	GTAACCTTCTTCATCGTATTCTCTCTACTTACCGTGTAGAAAGCTTCTCTAGTCAAGAGAC	58756

Qy 485 AspaSpGlnhapGluIleSerAlaLeuValSerSerTyrIleuGlyProSerThrSerPhe 504
Db 58757 GATGACCAAGATGAATATCATGCTCTTGTGAGTAGTACTCGCGCGTCAACATCTTTT 58816
Qy 505 ProHisArgSerArgArgSerProGluTyrMetValProLeuProHisGlyGlyLeuGly 524
Db 58817 CCTCATCGCTCAAGAAGATCCCGGAATATATGTTCCATCTCCACATGTTGGTTAGGA 58876
Qy 525 ArgSerValTyrAlaTyrGluHisLeuAlaProHisSerTyrSerProGlyHisGlyHis 544
Db 58877 AGAAGTATATGATATGATGAACATCTCGCCCCAAATCATATTTCCAGATCCAGACAT 58936
Qy 545 ArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArgHisProLeuGlnTyr 564
Db 58937 AGATTTTATCGACAGTACTCTCGCTCTTGGTTCCAGGACAGACATCCACTACAGTAC 58996
Qy 565 SerProProfileHisGlyGlnGlnGlnLeuProTyrGlyGlyGlnArgValTyrArgHis 584
Db 58997 TCTCTCTCAATTCTATGACACAAACAGTTACCATATGTTATACAAAGGTTTACAGACAT 59056
Qy 585 SerProSerGluGluArgTyrLeuGlyLeuSerAsnGlnArgSerProArgSerAsnSer 604
Db 59057 TCACATCTGAAAGAAATATTTGGGTTTATCCATCAAAAGTCTCTCCAGTAACTCA 59116
Qy 605 SerLeuAspProLys 609
Db 59117 TCATTAGACCCCAAA 59131
RESULT 11
ATCHRIV2 197975 bp DNA linear PLN 16-MAR-2000
LOCUS Arabidopsis thaliana DNA chromosome 4, contig fragment No. 2.
DEFINITION AL161472
ACCESSION AL161472
VERSION AL161472.2 GI:7267383
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 197975)
AUTHORS Lamar, B., Stoneking, T., Stumpf, J., Mewes, H.W., Lemcke, K. and
Mayer, K.F.X.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 197975)
AUTHORS EU Arabidopsis sequencing project.
TITLE Direct Submission
JOURNAL Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
lencke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de, Project
Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
E-mail: michael.bevan@bbsrc.ac.uk
COMMENT Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>
this fragment has an overlap with ATCHRIV1 at the 5' end and an
overlap with ATCHRIV3 at the 3' end.
FEATURES
source
1..197975
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
10352..17547
/gene="AT4G00450"
join(10352..10384,10432..10545,10605..10709,10849..11018,
11079..11423,11475..13674,13735..13823,13921..16530,
16677..16862,16976..17042,17092..17547)
/gene="AT4G00450"
/note="Contains Aminoacyl-transfer RNA synthetases
class-II signatures AA1343-1352; Prokaryotic membrane

lipoprotein lipid attachment site AA349-359
contains Ser 35:AI99580.1"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB80854.1"
/db_xref="GI:7267384"
/db_xref="SPTREMBL:O23057"
/translation="MLSPLFVISIGAVKCLRAINEGRALKRKVKIKRLSDSNPFQIF
CAEPAGGVYGLSGSLCKPKGPPEORSCGEETKRWIESLSQHKELRLANIPG
YRKILFVLRNNVPLLRATWFIKYVILNQHVCNAINWCDLVLLFLEOHLTHKLL
GVNSGMLNLCNTSMNFCGHIAHFLSKLEIGHRCFTQDOCKRIVQHQAFAR
KHLTYLNGIWCVFYGLMGFFLLSLIGPSSYSVNFHFALQKEIFELQL
LLPIYGVLESIVLSQTVSLVAIVFIOEPAPGSDLVNRRRAFTLLEMLV
VLLAAPPTFVADFFPLPSPAGCNDSYTKAVENKELASNAEIAEQGGRG
VLSRPELSPDYTTIORSADDAKIASAGYPOHNAKVAQALDKAISDIDRAAYS
VLFEDLNGAVDEAWITDVSCLRSSLRWIGAITSFVCSVFLEIENATCDPRFRAG
VFKDIFSGRCKDCQVILVQLKQILGGEFTARKGNCKNNELGUSKPSGSDAFE
SPGPHLDIIVICDQHEVHGKAKRLQLLFEIIRSGIFNFIATVRIQIVSGMDVIQ
PAVDPERMRHRIILKQLPGCFVHETLEEALFGDGLSEAVRTYSNRRLLRELLV
EKGYNNLVLSQSKKISTSLSVIFPRACNAKSNKSPRKHTKSVDIRELKRI
SALLQPPQMSCGVETPDEFQNSVGRSSGVSKMDQPEATPGCEDCRARPRMND
EKSCYQNSPIASDEEDNWKIKGSKTVESLKVDFQIETKQVPRGRQWARKTQ
LAOLQAARIEGSGASTSHVCDNKVSCPHGPGVGEKGVVDVFTSTPYDMYSVGN
SLKQLQFVKRSIAWLTAVRLVLEEPQKSSVRVGONRGAPVEEKNTIEMKLGADE
LSRATKIYAAVSALVSGITHGVIITLRIVGLLKLKDYLDVQVFRVTKSSNSGARS
MGASKVSPFIEYVHWFLVGNCKTVSEGLVLEIVGESSVVAISRMGMPLKLVFP
PAYSIATVLRFPVSNNSNVSHEDTHRLYQSLTMAFHDVIKHLPRDPCFRFTQ
LYELIADSTDAEFASFESHLGMLKSAFAPLARLFLNLDICKVPSGSHGEG
VSEAKNRHQNGTKLVKLVSLDCLQPAKEHWQVLELRLINLEQALAEKLENDHPL
TDAIRSCTSEKPDASEKNFIQLLTRLVDPADVPLFSEVHLFGRSVEDSMLK
QAEVLAGDVLFGKRTIQLIIVGESKGLPTKPOFKPMGNCNSSSDHTANKAG
KXREFTSIIEGEVIEGSGSRKVLPRVLDENSPVGYGITTERARVQLVLPIDQ
SDSEFTSVNELVRQFNEISQSSVTRNTTNKMGTSAGEISEISGSPKTKGL
RGSPSLARRSANTDTSPFPAALRASMLRLOFLRLLLPVCIGSPFNKTHAL
ASTIVLLGSRVYEDYAVCPRSBELSKAETSTDFSSMADLSSEVLFDRREITGLS
TNFKSLKLVCLLGCKNELSRQLPDTRWRIOQAAMPILLPSRLGSLSCQPHSVPTTAL
TLVQPSGTAAGTNQRNSPAISKSGTAAAGKLPMTLLAPHQOQEAONTVDVDPWT
LEDGTSGLSSSNASNSDMMANLRATCWLKAVRVRTDITLYVSGSDDS"
10352..10384
/gene="AT4G00450"
/number=1
10385..10431
/gene="AT4G00450"
/number=1
10432..10545
/gene="AT4G00450"
/number=2
10546..10604
/gene="AT4G00450"
/number=2
10605..10709
/gene="AT4G00450"
/number=3
10710..10848
/gene="AT4G00450"
/number=3
10849..11018
/gene="AT4G00450"
/number=4
11019..11078
/gene="AT4G00450"
/number=4
11079..11423
/gene="AT4G00450"
/number=5
11424..11474
/gene="AT4G00450"


```
QY 261 GlyLeuAlaAlaGlyLeuMetAspAlaArgGlyLeuLeuLeuValAlaCysPhe 280
DB 76493 GGATAGCTGGCGTGCAGAAATGCATGCAAGGGGTTTGCTTTTACTAGTTGCTTGTTT 76552
QY 281 GlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuLeuLeuMetSerGlySer 300
DB 76553 GGTGTTCTCTCAAACTTTAGGAGTACAGATTGCTGGATTGTGTAAGAGTGGTTCG 76612
QY 301 AsnGluLeuAlaGlyAlaLeuLysArgSerGlnPheLeuValProMet----- 316
DB 76613 AATGAGATTGCGGTGCTTTGAACGGTCCACAGTTTCAGGTACCATA-TTCTGTTCTCAC 76671
QY 316 ----- 316
DB 76672 TCGGTGAATTCATTGCAAGGTGGTCTCTTTGTTGACATCATCGACCAACATCAAGTT 76731
QY 316 ----- 316
DB 76732 CCATCTTTTGTTCGATAAGCTTGATGGTATAAACTAGGAGAGACATCAAAATATTAG 76791
QY 316 ----- 316
DB 76792 AGTCAATGACTGATTGAGCCAAATCCTAGTAGAAATTAATCTGGAAGAACTTGAAC 76851
QY 316 ----- 316
DB 76852 TCTCAACCATAGTTTGTGACGAAATTTGTTCTTGTGAGAACCAAAATGATAGGCTATTG 76911
QY 316 ----- 316
DB 76912 CTTGAATAGTGTTCTTGTTGGTTTCCAATATTGGAAGTTAAATCATATGACTTAGCT 76971
QY 316 ----- 316
DB 76972 GTTGGATACTAATTAAGCTTAAGCAATGCCAACTCTAAGAGTGCTACTTACACAATATT 77031
QY 317 -----ValSerGlyLeuValGluSerSerIleLysArgGlyMetHisIleGluAlaLeu 334
DB 77032 CTATTGGTCATAGGTATAGTTGAATCAAGTATCAAGGTGAATGCTATATTGAAGCTCTT 77091
QY 335 GluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSer 354
DB 77092 GAAATGGTTTATACCTTTGGCATGGAGATAAGTTTTCAGCTGCTCTAGTTCTAACTTCA 77151
QY 355 PheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeu 374
DB 77152 TTCTTAAGATGAGCAGGAGTCTATTGAGAGGGCAAAACGGAAAGCCAGTCAACCGCTG 77211
QY 375 AlaPheLys----- 377
DB 77212 GCATTTGT-ATGAACCTTCCCTTGCACATTTATGTACCTTTATGAATCTTTATCATCAT 77270
QY 378 -----GluAlaAlaThrLysGlnLeu 384
DB 77271 CTGAGTCTGACCAATTGATATATTATTCTCAACAGAAAGAGCGGCTCAAAAGCAGCTA 77330
QY 385 AlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeuAspProAlaLysGlu 404
DB 77331 GCTGTGTTATCATCAGTTTATGAGTATGAGAGACTCAAGTTAGTATCTTCGCGAAGAA 77390
QY 405 LeuProGlyTyrGlnIleLysGluGlnIleValSerLeuGluLysAspThrLeuGlnLeu 424
DB 77391 CTACCAGGATGGCAGATCAAAAGAGCAAAATTTGTAGCTTGGAGAAAGACACTCTTCAGCTC 77450
QY 425 AspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGluGluAlaAlaLeuAla 444
DB 77451 GACAAAGAGATGGAAGAAAGCAAGATCTCTCAGTTTAAATGGAGGAGCGGCACCTTGCC 77510
QY 445 LysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSerProMetGluMetProPro 464
DB 77511 AAGAGAAATGATAAACACAGATAAAAGCTCAAAGTTATCAACCATGGAAATGCCACCA 77570
```

```
465 ValThrSerSerSerTyrSerProIleTyrArgAspArgSerPheProSerGlnArgasp 484
77571 GTAATCTTTCATCGTATTCTCTTATCTACCGTGATAGAAAGCTTTCTTAGTCAAAGAGAC 77630
485 AspAspGlnAspGluIleSerAlaLeuValSerSerTyrLeuGlyProSerThrSerPhe 504
77631 GATGACCAAGATGAAATATCAGCTCTTGAGTAGTTACCTCGGCCCTGCAACATCTTTT 77690
505 ProHisArgSerArgArgSerProGluTyrMetValProLeuProHisGlyLysGlyLeuGly 524
77691 CTTCTATCGTCAAGAGATCCCGGAATATATGTTTCCACTTCCACATGGTGGGTAGGA 77750
525 ArgSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSerProGlyHisGlyHis 544
77751 AGAAGTGATATGATGATGATGAAATCTGGCCCAAAATTCATATCTTCCAGGTCAAGGACAT 77810
545 ArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArgHisProLeuGlnTyr 564
77811 AGACTTCATCGACAGTACTCTCCGCTTTGGTTTCAGGACAGACATCCACTACAGTAC 77870
565 SerProIleHisGlyGlnGlnLeuProTyrGlyIleGlnArgValTyrArgHis 584
77871 TCTCTCCCAATTCATGACCAACACAGATTACCATATGTTATACAAAGGTTTACAGACAT 77930
585 SerProSerGluGluArgTyrLeuGlyLeuSerAsnGlnArgSerProArgSerAsnSer 604
77931 TCACCATCTGAAAGAAAGATATTGGGTTTATCCAAATCAAGAGTCTCTCGCAGTAACTCA 77990
605 SerLeuAspProLys 609
77991 TCATTAGACCCCAAA 78005
```

RESULT 12
AY092681
LOCUS
DEFINITION
Arabidopsis thaliana-cultivar RSCH-4 FRIGIDA protein gene, partial cds.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 558)
Hagenblad, J. and Nordborg, M.
Sequence Variation and Haplotype Structure Surrounding the
Flowering Time Locus FRI in Arabidopsis thaliana
Genetics 161 (1), 289-298 (2002)
JOURNAL
MEDLINE
PUBMED
2 (bases 1 to 558)
Hagenblad, J. and Nordborg, M.
Direct Submission
Submitted (27-MAR-2002) Department of Genetics, Lund University,
Solvegatan 29, Lund SE-223 62, Sweden
JOURNAL
FEATURES
source
1..558
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="RSCH-4"
/db_xref="taxon:3702"
<1..>558
/product="FRIGIDA protein"
<1..>558
/codon_start=1
/product="FRIGIDA protein"
/protein_id="AA12134.1"
/db_xref="GI:"AA12134.1"
/translation="KIVETESMDITIGSQSPQLKSIDELAAFSVAVETFKRQFD
DIOKHIESIENAIKSLNSGVVLAARNNPHQPNVSVETTVTSQSQEI
VPETSNKPGGRMCMCSKGLRKYIYANISDQAKLMEEIPSAKLAKPAKFLDCI

Db	241	TGCGCTCCGGGAACAATGTATCTGTAGAAACCAACCGTACTGTGTAGGCCAACCGTCTCAG	300
QY	134	GlulleValProGluThrSerAsnLysProGluGlycylArgMetCysGluLeuMetCys	153
Db	301	GAGATTGTACCGGAGAGCGTCGATATAAACCAGGAGGGGACGTATGTGTGAGTTGATGTGT	360
QY	154	SerLysGlyLeuAlaGlyLeuTyrlleTyAlaasnIleSerAspGlnAlaLysLeuMetGlu	173
Db	361	AGCAAAGGTCTGCCTTAATCATATACATATACCGAATATCTTGATCAAGCTAAGTTAATGGAA	420
QY	174	GlulleProSerAlaLeuLysLeuAlaLysGluProAlaLysPheValLeuAsePysille	193
Db	421	GAGATTCTCTTTCAGCTTTCGAAATTCGCCAAGGAGCGACGGAAGTTGTGTATGGATTGTA	480
QY	194	GlyLysPheTyLeuGlnGlyArgAlaPheThrylsGluSerProMetSerSerAla	213
Db	481	GCGAAGTTTTTACTTTACAAGGCGTAGAGCATTACTTAAGAAGTCGGCTATGAGCTCTGG	540
QY	214	ArgGlnValSerLeuLeu 219	
Db	541	AGACAAGTTTCGCTTCIT 558	
RESULT 15			
AY092693			
LOCUS	Arabidopsis thaliana cultivar Ler FRIGIDA protein gene, partial cds.	558 bp DNA linear	PLN 03-JUN-2002
DEFINITION	Arabidopsis thaliana cultivar Ler FRIGIDA protein gene, partial cds.		
ACCESSION	AY092693		
VERSION	AY092693.1	GI:20160140	
KEYWORDS	Arabidopsis thaliana (thale cress)		
SOURCE	Arabidopsis thaliana		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi		
REFERENCE	1 (bases 1 to 558)		
AUTHORS	Hagenblad,J. and Nordborg,M.		
TITLE	Sequence Variation and Haplotype Structure Surrounding the Flowering Time Locus FRI in Arabidopsis thaliana		
JOURNAL	Genetics 161 (1), 289-298 (2002)		
MEDLINE	22013871		
PUBMED	12019242		
REFERENCE	2 (bases 1 to 558)		
AUTHORS	Hagenblad,J. and Nordborg,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-MAR-2002) Department of Genetics, Lund University, Solvegatan 29, Lund SE-223 62, Sweden		
FEATURES	Location/Qualifiers		
source	1..558		
mRNA	/organism="Arabidopsis thaliana"		
CDS	/mol_type="genomic DNA"		
	/cultivar="TSU-0"		
	<1..>558		
	/product="FRIGIDA protein"		
	<1..>558		
	/codon_start=1		
	/product="FRIGIDA protein"		
	/protein_id="AA12144.1"		
	/db_xref="GI:20160137"		
	DIOKHIESINAIDSKLSNGVLLAARNNHOPMLSPNNVSVEITTVTSQPSDEI		
	VPTSNKPGGRMCELMCKGLRKYIANISDQAKLMEEIPSAKLAKPAKFVLDCI		
	GKFLQGRRAFTKESPMSARQNSLL"		
ORIGIN			
Alignment Scores:			
Pred. No.:	8.2e-53	Length:	558
Score:	931.00	Matches:	185
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.46%	Mismatches:	0
Query Match:	29.99%	Indels:	0
DB:	8	Gaps:	0
US-09-890-475-1 (1-609) x AY092691 (1-558)			
QY	34	LyeileValGluThrGluSerThrSerMetAspIleThrIleGlyGlnSerLysGlnPro	53
Db	1	AAGATTGTCCAACAGAGCTCTACAGTATGGACATTAGCATCGGTCAATCTAAGCAGCT	60
QY	54	GlnPheLeuLysSerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLys	73
Db	61	CAATTTTGAATCCATAGCAGNAATTAGTCGGTTTCAGTTGCGATGGAACATTCAAA	120
QY	74	ArgGlnPheAspAspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLys	93
Db	121	CGCAATTTCGATGATATCAGAGCACATCGATCGAATCGAAACGCAATTGATTCCAA	180
QY	94	LeuGluSerAsnGlyValValLeuAlaAlaAcGAsnAsnAsnPheHiscLnProMetLeu	113
Db	181	CTCAGAGTAGTAAACGGCTGTTCCTCGCGCGCGGAGAACATATTTTCATCAGCCGATGTA	240
QY	114	SerProProArgAsnValSerValGluThrValThrValSerGlnProSerGln	133

US-09-890-475-1 (1-609) x AY092693 (1-558)

QY 34 LysIleValGluThrGluSerThrSerMetAspIleThrIleGlyGlnSerLysGlnPro 53
 Db 1 AAGATTGTCGAAACAGAGCTTACAGATGATGACATTAACGATCGGTCAATCTAAGCAGCCT 60

QY 54 GlnPheLeuLysSerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLys 73
 Db 61 CAATTTTGAATTCCTAGACGAATTAGTCGCTTTCAGTTGAGTGGAAACATTCCAAA 120

QY 74 ArgGlnPheAspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLys 93
 Db 121 CGCCATTTCGATGATATTAGAGCAGCATCGAGTCAATCGAAACGCAATTGATTCCAAA 180

QY 94 LeuGluSerAsnGlyValIleValAlaAlaAlaArgAsnAsnAsnPheHisGlnProMetLeu 113
 Db 181 CTCGAGAGTACGGGGTGTCTCTCCCGCGGGACAAATAATTCATCAGCCGATGTTA 240

QY 114 SerProProArgAsnAsnValSerValGluThrThrValThrValSerGlnProSerGln 133
 Db 241 TCGCCTCCGCGAACAATGTATCTGTAGAAACCCGTCCTCTGAGCCACCGTCTCAG 300

QY 134 GluIleValProGluThrSerAsnLysProGluGlyGlyArgMetCysGluLeuMetCys 153
 Db 301 GAGATTGTACCGGAGACGTGGAATAACCGAGGGGGACGTATGTGTGAGTTGATGTGT 360

QY 154 SerLysGlyLeuArgLysTrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGlu 173
 Db 361 AGCAAGGTCTCGGTAAATACATATACCGGAATATCTCTGATCAAGCTAAGTTAATGGAA 420

QY 174 GluIleProSerAlaLeuLysLeuAlaLysGluProAlaLysPheValLeuAspCysIle 193
 Db 421 GAGATTCCTTCAGCTTTGAAATTGGCCCAAGGAGCCAGCGAAGTTGTATTGCGATTGTATT 480

QY 194 GlyLysPheTyrLeuGlnGlyArgAlaPheThrLysGluSerProMetSerSerAla 213
 Db 481 GGCAAGTTTACTTACAGGGGGGTAGAGCATTTACTAAGAGTCGCCCTATCAGCTCTCGG 540

QY 214 ArgGlnValSerLeuLeu 219
 Db 541 AGACAAAGTTTCGCTTCTT 558

Search completed: February 28, 2004, 04:51:32
 Job time : 5820 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 27, 2004, 21:18:27 ; Search time 574 Seconds
(without alignments)
4507.235 Million cell updates/sec

Title: US-09-890-475-1

Perfect score: 3104

Sequence: 1 MSNPPTVAAQPTTIANPLL.....RYLGLSNQRSPRNSSLDPK 609

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747736

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n_model -DEV=xlp
-Q/cn2_1/USPTO_spool_p/US09890475/runat_24022004_135311_5281/app_query.fasta_1.775
-DB=N Geneseq_25Jan04 -QFMI=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09890475 @CGN 1 1.885 @runat_24022004_135311_5281 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04.*

1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001as.*
5: Geneseq2001bs.*
6: Geneseq2002s.*
7: Geneseq2003as.*
8: Geneseq2003bs.*
9: Geneseq2003cs.*
10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3104	100.0	2257	AAA63669	AAA63669 cDNA sequ
2	2981	96.0	3761	AAA63668	AAA63668 H51 genom
3	349	11.2	2155	AAC33344	AAC33344 Arabidops
4	340	11.0	2158	AAC46135	AAC46135 Arabidops
5	337	10.9	1632	ADC03365	ADC03365 Rice flow
6	329	10.6	1602	ADC03469	ADC03469 Maize flo
7	296.5	9.6	1964	AAC49876	AAC49876 Arabidops
8	296.5	9.6	1970	AAC35353	AAC35353 Arabidops

9	293	9.4	1377	9	ADC03367	ADC03367 Rice flow
10	288.5	9.3	1865	3	AAC38630	AAC38630 Arabidops
11	234.5	7.6	1524	6	ABZ13112	ABZ13112 Arabidops
12	224	7.2	947	9	ADC03447	ADC03447 Wheat flo
13	152	4.9	7785	4	ABL17653	ABL17653 Drosophil
14	151	4.9	3312	2	AAC32435	AAC32435 Rat TAO1
15	151	4.9	3312	7	ABT40852	ABT40852 Toxicity
16	148.5	4.8	3046	8	AAL62252	AAL62252 Human TAO
17	148.5	4.8	3824	2	AAC40489	AAC40489 Human SUL
18	148.5	4.8	4535	8	AAL62250	AAL62250 Human TAO
19	148.5	4.8	4536	8	AAL62249	AAL62249 Human TAO
20	148.5	4.8	4620	7	ACC50205	ACC50205 Breast ca
21	148.5	4.8	4620	8	AAL62248	AAL62248 Human TAO
22	148.5	4.8	4972	6	ABN59739	ABN59739 Novel hum
23	145.5	4.7	2760	6	ABK32901	ABK32901 DNA encod
24	145	4.7	2187	6	ABS51736	ABS51736 Human nov
25	145	4.7	3045	6	AAS94957	AAS94957 Human DNA
26	144.5	4.7	3003	8	AAL62251	AAL62251 Human TAO
27	144	4.6	2684	6	ABS51734	ABS51734 Human nov
28	144	4.6	2708	9	ADA08249	ADA08249 DNA encod
29	144	4.6	2756	6	ABK70288	ABK70288 Human lun
30	144	4.6	2756	9	ADA08248	ADA08248 DNA encod
31	143	4.6	3270	7	ABT16471	ABT16471 Human int
32	143	4.6	18506	4	ABL05592	ABL05592 Drosophil
33	142	4.6	3114	2	AAV02800	AAV02800 Human RHA
34	142	4.6	3114	6	AAS17496	AAS17496 Human cdn
35	141.5	4.6	4772	4	ABL03209	ABL03209 Drosophil
36	141	4.5	2020	6	ABS51735	ABS51735 Human nov
37	140.5	4.5	2511	4	AAS1630	AAS1630 Human pol
38	140.5	4.5	2609	9	ADA08250	ADA08250 DNA encod
39	140.5	4.5	3587	9	ADB63157	ADB63157 Human cdn
40	140	4.5	3718	6	ABK63564	ABK63564 Rat sequ
41	137.5	4.4	6558	3	AAC45602	AAC45602 cDNA sequ
42	136	4.4	6455	6	ABK84151	ABK84151 Human cdn
43	136	4.4	6455	9	ACF79923	ACF79923 Breast ca
44	135	4.4	7080	5	AAS85528	AAS85528 DNA encod
45	135.5	4.4	14278	4	ABL06088	ABL06088 Drosophil

ALIGNMENTS

RESULT 1
AAA63669
ID AAA63669 standard; cDNA; 2257 BP.
XX AC AAA63669;
XX AC
XX AC
DT 04-DEC-2000 (first entry)
XX AC
DE cDNA sequence of the FRI gene of Arabidopsis.
XX H51; one locus-FRIGIDA; FRI gene; flowering time; blotting;
XX flower initiation; stem elongation; flower production; ss.
XX Arabidopsis sp.
XX Key Location/Qualifiers
XX CDS 362..2191
XX FT /*tag= a
XX FT /product= "polypeptide which alters flowering time"
XX PN WO200046358-A2.
XX PD 10-AUG-2000.
XX PF 25-JAN-2000; 2000WO-GB000197.
XX PR 05-FEB-1999; 99GB-00002660.
XX PA (PLAN-) PLANT BIOSCIENCE LTD.
XX PI Johanson U, West J, Dean C;
XX


```

QY 581 ValTyrArgHisSerProSerGluGluArgTyrLeuGlyLeuSerAsnGlnArgSerPro 600
AAA63668
DB 2102 GTTTACAGACATTACCATCTGGAAGAAATATTTGGTTTATCCAAATCAAAGGTCTCTCT 2161
QY 601 ArgSerAsnSerSerLeuAspProIys 609
DB 2162 CGCAGTAACATCATATTAGACCCCAAA 2188

RESULT 2
ID AAA63668 standard; DNA; 3761 BP.
XX AAA63668;
DT 04-DEC-2000 (first entry)
DE H51 genomic sequence of first 17 kb of 84M13 encompassing FRI gene.
KW H51; one locus-FRIGIDA; FRI gene; flowering time; blotting;
KW flower initiation; stem elongation; flower production; ss.
XX Arabidopsis sp.
XX WO200046358-A2.
XX 10-AUG-2000.
XX 25-JAN-2000; 2000WO-GB000197.
XX 05-FEB-1999; 99GB-00002660.
XX (PLAN-) PLANT BIOSCIENCE LTD.
XX Johanson U, West J, Dean C;
XX WPI; 2000-532899/48.
XX New nucleic acid derived from the FRI locus of a plant, e.g. Arabidopsis,
PT encoding a polypeptide capable of specifically altering the flowering
PT time of a plant.
XX Claim 4; Fig 4; 73pp; English.
XX The present sequence represents the (late flowering) H51 genomic sequence
CC of the first 17 kb of cosmid 84M13 encompassing FRI (one locus-FRIGIDA)
CC locus of Arabidopsis. The FRI gene encodes a polypeptide capable of
CC specifically altering the flowering time of a plant. The FRI
CC polynucleotide is used to transform plants, so that the flowering time of
CC a plant is altered. This is used, for example, for plants in which the
CC leaves or tubers are a commercial product, where it is desirable to avoid
CC 'blotting' (initiation of flowers and stem elongation) at too early a
CC stage. Conversely, it may be desirable to alter flowering under certain
CC circumstances e.g. to vary flower production across the seasons
XX
SQ Sequence 3761 BP; 1174 A; 717 C; 775 G; 1095 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6.4e-232 Length: 3761
Score: 2981.00 Matches: 608
Percent Similarity: 78.96% Conservative: 0
Best Local Similarity: 78.96% Mismatches: 1
Query Match: 96.04% Indels: 162
DB: 3 Gaps: 2

US-09-890-475-1 (1-609) x AAA63668 (1-3761)
QY 1 MetSerAsnTyrProProThrValAlaAlaGlnProThrThrAlaAsnProLeuLeu 20
DB 574 ATGTCCAATTATCCACCGAGCGTGGCGGCAACCCACACGACGGCGAATCCACTGCTG 633
QY 21 GlnArgHisGlnSerGluGluArgGluLeuProIysIleValGluThrGluSer 40

```

```

DB 634 CAGCGACATCAATCTGAACAGCGACGAAGAGAATTTACCGAAGATTGTCGAACACAGACTCT 693
QY 41 ThrSerMetAspIleThrIleGlyGlnSerIysGlnProGlnPheLeuIysSerIleAsp 60
DB 694 ACAAGTATGGACATTACGATCGGTCAATCTAAGCAGCGCTCAATTTTGAATCCATAGAC 753
QY 61 GluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAspAspLeuGln 80
DB 754 GAATTAGCTGCGTTTTTCAGTTGCAATGGAAACATTCMAACGCCAATTTCGATGATCTTCAG 813
QY 81 LysHisIleGluSerIleGluAsnAlaIleAspSerIysLeuGluSerAsnGlyValVal 100
DB 814 AAGCACATCGAGTCAATCGAAACGCAATTGATTCGAACTCGAGATAACGCGCTGTC 873
QY 101 LeuAlaAlaArgAsnAsnPheHisGlnProMetLeuSerProProArgAsnVal 120
DB 874 CTCGCCGCGGCAACAATAATTCATCAGCGCATGTTATCCCTCCCGGGAACATGTA 933
QY 121 SerValGluThrThrValThrValSerGlnProSerGlnGluIleValProGluThrSer 140
DB 934 TCTGTAGAAACCCACCGTCACTGTGAGCAACCGTCTCAGGAGATTGTACCGGAGACGTG 993
QY 141 AsnLysProGluGlyArgMetCysGluLeuMetCysSerIysGlyLeuArgIysTyr 160
DB 994 AATAAACCGGAGGGGACGTATGTGTGATTTGATGTAGCAAAAGGTCTGCGTAATAC 1053
QY 161 IleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAlaLeuLys 180
DB 1054 ATATACCGAATATCTCTGATCAAGCTAAGTTAATGGAAGAGATTCTCTCAGCTTTGAA 1113
QY 181 LeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGly 200
DB 1114 TTGGCCAGGAGCCAGCGAAGTTTGTATTGGATTGTATTGGCAAGTTTACTTACAAGGG 1173
QY 201 ArgArgAlaPheThrIysGluSerProMetSerSerAlaArgGlnValSerLeuLeuIle 220
DB 1174 CGTAGAGCATTTACTAAAGAGTCGCTATGAGCTTCGAGACACAGTTTCGGTCTTATA 1233
QY 221 LeuGluSerPheLeuLeuMetProAspArgGlyLysGlyLysValIysIleGluSerTyr 240
DB 1234 CTGGAGTCTTTTCTTAATGCTGATCGTGTAAAGGGAAGGTGAAGATTGAGAGTTGG 1293
QY 241 IleLysAspGluAlaGluThrAlaAlaValAlaIleTyrArgLysArgLeuMetThrGluGly 260
DB 1294 ATTAAAGATGAGCGGAGACGCTGCTGTGTGTGAGGAAAGGTGTGATGACTGAAGA 1353
QY 261 GlyLeuAlaAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPhe 280
DB 1354 GGATTAGCTGCGGCTGAGAAATGGATGCAAGGGTTTCTTTTACTAGTTGCTTTT 1413
QY 281 GlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySer 300
DB 1414 GGTGTTTCCCTTCAAACTTTAGGAGTACAGATTTGCTGGATTTCATAGGATGAGTGTG 1473
QY 301 AsnGluIleAlaGluAlaLeuLysArgSerGlnPheLeuValProMetValSer 318
DB 1474 AATGAGATTGCGGCTGCTTTGAGCGGTCAAGTTTCTTGTCCCTATGCTCTCAGGTACC 1533
QY 318 ----- 318
DB 1534 ATATTCTGTCTCACTCGGTGAATTTCAATTGCAAGGTGTTCTCTTTTGTGACATCATC 1593
QY 318 ----- 318
DB 1594 GACCAACATCAAGTTCCTATCTTTGTTTTCGATAAGCTTGATGGTATAAACTAGGAGAC 1653
QY 318 ----- 318
DB 1654 ACATCAATATTATTAGAGTGCATGACTGATTCAGCCAAATCCTAGCTAGAGAAATTAATCTG 1713
QY 318 ----- 318
DB 1714 GAAAGAACTGGAACTCTCAACCATAGGTTTGTGTACGAAATTTGTTGTTGTTCAGAACCA 1773

```



```
QY 318 ----- 318
DB 1774 AATGATAGGCTATTGCTTGAATAGTGTCTTGTGTTTCCAAATATTGGAAGTTAAAA 1833
QY 318 ----- 318
DB 1834 TCGTATGACTTAGCTGTGGATCTAAATTAAGCTTAAGCAATGCCAACTCTAAGAACTGG 1893
QY 319 ----- 319
DB 1894 TACTTACACAATATCTATTGCTCATAGTATAGTGAATCAAGTATCAAGCGTGGAAATG 1953
QY 330 HisLeGluAlaLeuGluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAla 349
DB 1954 CATATTGAAGCTCTTGAGATGTTTATACCTTTGGCATGGAGGATAAGTTTTCAGCTGCT 2013
QY 350 LeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLys 369
DB 2014 CTAGTTCTTAACCTTCTTAAGATGAGCAAGGAGTCACTTGAGAGGGCAAAACGGAA 2073
QY 370 AlaGlnSerProLeuAlaPheLys ----- 377
DB 2074 GCCCAGTCACCGCTGCCATTGT-ATGAACCCCTTCCCTTGCACTTATGTACCTTTTATGA 2132
QY 378 ----- GluAla 379
DB 2133 ACTCTTTATCATCTGAGCTGACCATGATGATATATTTTCTCAACAAAGAGCG 2192
QY 380 AlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeu 399
DB 2193 GCTACAAAGCAGCTAGCTGTGTATCATCAGTATGAGTGTGTGAGACTCACAAGTTA 2252
QY 400 AspProAlaLysGluLeuProGlyTyrGlnLysGluGlnLeuValSerLeuGluLys 419
DB 2253 GATCTGCGAAGAACTACAGAGTGGCAGATCAAGAGCAAAATGTTAGCTTGGAGAA 2312
QY 420 AspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGlu 439
DB 2313 GACACTCTTCAGCTCGCAAGAGATGGAAGAGAAAGCAAGATCTCTCAGTTTAAATGGAG 2372
QY 440 GluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnLysArgProArgLeuSerPro 459
DB 2373 GAAGCGGCATCTGCCAAGAGATGATATACCAACAGATAAAACGTCACAGTTGTCACCC 2432
QY 460 MetGluMetProProValThrSerSerSerTyrSerProLysArgAspArgSerPhe 479
DB 2433 ATGGAAATGCCACAGTAATCTTCTCATGCTATTTCTCTATCTACCGTGATAGAGCTTT 2492
QY 480 ProSerGlnArgAspAspGlnAspGluLysSerAlaLeuValSerSerTyrLeuGly 499
DB 2493 CCTAGTCAAAAGAGACGATGACCAAGATGAAATATCAGCTCTTGTGAGTAGTTACCTCGGC 2552
QY 500 ProSerThrSerPheProHisArgSerArgSerProGluTyrMetValProLeuPro 519
DB 2553 CCGTCAACATCTTTTCTCATGCTCAAGAGATCCCGGAATATATGTTCCACTTCCA 2612
QY 520 HisGlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSer 539
DB 2613 CATGGTGGGTAGGAAGAAGTGTATATGATATGAACATCTGGCCCCCAAAATTCATCTCT 2672
QY 540 ProGlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArg 559
DB 2673 CCAGGTCCAGGACATAGATCTCATGACAGTACTCTCCGCTCTTGGTCCAGGACAGAGA 2732
QY 560 HisProLeuGlnTyrSerProProLysHisGlyGlnGlnLeuProTyrGlyIleGln 579
DB 2733 CATCCACTACAGTACTCTCTCCAAATTCATGACACACAGTACCATATGCTATACAA 2792
QY 580 ArgValTyrArgHisSerProSerGluGluArgTyrLeuGlyLeuSerAsnGlnArgSer 599
DB 2793 AGGGTTTACAGACATTCACCATCTGAAGAAGATATTTGGGTTTATCCCAATCAAAGGTCT 2852
```

600 ProArgSerAsnSerSerLeuAspProLys 609
2853 CCTCGCAGTAACCTATCATTAGACCCAAA 2882

RESULT 3
AAC33344
ID AAC33344 standard; DNA; 2155 BP.
AC AAC33344;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 2690.
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 03-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 07-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.

392	Db	GCATTTTCATGGCGCTTGAGAGGCTTTTGAAGAGAGAGGTTTCATGAGCTTGAAAGATCAAGA	451
80	Qy	nLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsnGlyValVa	100
452	Db	AAAAGAGTATGAACCAACCAAAACAAGGAAAGCTCAAGAACTGTTGGAGAAAAAAGAGCAGC	511
100	Qy	LleuAlaAlaArgAsn	105
512	Db	TGTTGAGCTTAAGGAGAAAGCGCGCTTGAAGAGCTTCAGAAGAAGAGACACCACTAT	571
106	Qy	-----AsnAsnPheHisGlnProMetLeuSerPr	115
572	Db	GTITACTATAAACACAGCTGTTTGGATAAGTATAACAAC-----GCTCTGTCTAGTAAACC	625
115	Qy	o-----ProArgAsn-----AsnValSerValGl	123
626	Db	TTCTGTTGGTCAACGTTTGGCCACAAAATGCTGTAGAGATTCTCAATGCTTTGCTGTCG	685
123	Qy	uThrThrValThrValSerGlnPro-----SerGlnGluIleValPro--	137
686	Db	TGATAGTATTACCGATGATAATCCGGATGGTATTGTGCAAGATGTTTCAGATCTTCACCTGT	745
138	Qy	-----GluThrSerAsnLysProGluGlyGlyArgMetCysGluLeuMetCy	153
746	Db	GATCGGAAATTAATCAGGTGAAGGCTTATCCACAGCTTACTGAACTGTGCGGGATATGGA	805
153	Qy	sSerLysGluLeuArgLysTyrlleTyrlleAlaAsnIleSerAspGlnAlaLysLeuMetGl	173
806	Db	CTCAACGGGCTCCATAGTATTGTATCAGATAACCGCAAGAACCTTGATCATCTAATAAGGA	865
173	Qy	uGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLysPheValLeuAspCysIl	193
866	Db	GGAGATTCTTATGCGTTCAGGGCTGCAGCAAAACCAGCCAGCTTAGTGTGGACTCTTT	925
193	Qy	eGlyLysPheTyr-----LeuGlnGlyArgArgAlaPheThrLysGl	207
926	Db	GGAAGGGTTTTACCTATGAGGCGCAACCTGCTGTATGGAAAA-----AAGA	973
207	Qy	uSerProMetSerSerAlaArgGlnValSerLeuLeuIleLeuGlu-----SerPheLe	225
974	Db	CGCTAACCTCTTGGCGATCGCAACCTGTATCATCTTGTATGAGAGTCCCTTAGCATACT	1033
225	Qy	uLeuMetPro---AspArgGlyLysGlyLysValLysIleGluSerTrpIleLysAspGl	244
1034	Db	GTTATCTGCTGAGCCGCACTCCCTTGCTGTTGTTCTCTCAAAAATGTTAAGCATAG	1093
244	Qy	uAlaGluThrAlaAlaValAlaTtpArgLysArgLeuMetThr---GluGlyGlyLeuAl	263
1094	Db	AGCCAAACTATTCCAGNAGATGGATCCACTCTGGAATCTCTTGACATGATGCTTG	1153
263	Qy	aAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPheGlyValPr	283
1154	Db	CAATGGTAATTCCTTTGGAGGCTCATGATTCCTCAACTGCTGCCACTTTTGTATTGT	1213
283	Qy	oSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySerAsnGluIl	303
1214	Db	TGCTGATTTTAAAGAGATGAATCTTTAAGCTGATCCCATGGTTCACGTCCGCTCA	1273
303	Qy	eAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIleValGluSe	323
1274	Db	GGCAGCTGAGCTCTCGCGTTTCCATTGGATTAGCTGAAAAAATGCTGTGTGATTGAAT	1333
323	Qy	rSerIleLysArgGlyMetHisIleGluAlaLeuGluMetValTyrThrPheGlyMetGl	343
1334	Db	TCTCGTGAACAGTGGAAACAGATTGATGCGGTTTAACCTGGCATTTCCGTTTGAACCTAC	1393
343	Qy	uAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLysGluSerPh	363
1394	Db	AGAACGTTCTCACCTGCTCTATTACTGAATCTTACTTGATAGAGCAAGCGCATCTTC	1453
363	Qy	eGluArgAlaLysArgLysAlaGlnSerProLeuAlaPheLysGluAlaAlaThrLysGl	383
1454	Db	CCCCCAAGCGACCTCGCAATGATCTCTCTGTTTCCAGATGATGTTCAATGAGCGGGA	1513

Qy	383	nLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeuAspProAlaLys	403
Db	1514	GCATTATAGGCCTCAAAACCGTTATAAAGTGATTCGAAGACATAGCCTGGAA-----GA	1567
Qy	403	sGluLeuProGlyTrpGlnIleLysGluGlnIleValSerLeuGluLysAspThrLeuGlu	423
Db	1568	GCAGTATCCCA-----GTTGAGCCACTTCACAAACGGATTCTCCA	1606
Qy	423	nLeuAspLysGluMetGluGluLysAlaArgSerLeuMetGluGluAlaLys	443
Db	1607	GCTTCGAGAAAGCCAAACGGACAAAAGAGAGCTACAGAACCCATGAAA-----	1655
Qy	443	uAlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSerProMetGluMetPr	463
Db	1656	-----CCTCAGCCAAAACGACACCGTGTGCT-----CAACC	1687
Qy	463	oProValThrSerSerSer-----TyrSerProIleTyr	474
Db	1688	CCGAGTCACATGACCAACAACAACATCAACAACAACAGACAGGATACGGTAGAGTCAT	1747
Qy	474	rArgAspArgSerPheProSerGlnArgAspAspGlnAspGluLeuSerAlaLeuVa	494
Db	1748	CCCTGAAAGG--TATCCGCAGTATGCTATGACACAGACCGCTCCTTAGCGGTCCAAT	1804
Qy	494	lSerSerTyrLeuGlyProSerThrSerPheProHisArgSerArgArgSerProGluTyr	514
Db	1805	CATGGCGACGACCAACCTCCACCACTCCTCCTCTCAGACTTACAGTTCAATCTCT-----	1859
Qy	514	rMetValProLeuProHisGlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAl	534
Db	1860	-----GCTCCT--GCACATCGGAACCTTCTACGCGAACTGCTACTCAGTACCAAGGCTCCTCC	1912
Qy	534	aProAsnSerTyr 538	
Db	1913	TCCTCCTCCTTAC 1925	
RESULT 4			
ID	AAC46135		
XX	AC	AAC46135 standard; DNA; 2158 BP.	
XX	AC	AAC46135;	
XX	DT	18-OCT-2000 (first entry)	
XX	DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 49031.	
XX	KW	Hybridisation assay; genetic mapping; gene expression control;	
XX	KW	protein identification; signal transduction pathway; metabolic pathway;	
XX	KW	promoter; termination sequence; ss.	
XX	OS	Arabidopsis thaliana.	
XX	PN	EP1033405-A2.	
XX	PD	06-SEP-2000.	
XX	PF	25-FEB-2000; 2000EP-00301439.	
XX	PF	25-FEB-1999; 99US-0121825P.	
PR	PR	05-MAR-1999; 99US-0123180P.	
PR	PR	09-MAR-1999; 99US-0123548P.	
PR	PR	23-MAR-1999; 99US-0125788P.	
PR	PR	25-MAR-1999; 99US-0126264P.	
PR	PR	29-MAR-1999; 99US-0126785P.	
PR	PR	01-APR-1999; 99US-0127462P.	
PR	PR	06-APR-1999; 99US-0128234P.	
PR	PR	08-APR-1999; 99US-0128714P.	
PR	PR	16-APR-1999; 99US-0129845P.	
PR	PR	19-APR-1999; 99US-0130077P.	
PR	PR	21-APR-1999; 99US-0130449P.	
PR	PR	23-APR-1999; 99US-0130510P.	
PR	PR	23-APR-1999; 99US-0130891P.	

PR	28-APR-1999;	99US-01311449P.	PR	21-JUL-1999;	99US-01448141P.
PR	30-APR-1999;	99US-01322048P.	PR	21-JUL-1999;	99US-01450866P.
PR	30-APR-1999;	99US-01322047P.	PR	21-JUL-1999;	99US-01450888P.
PR	04-MAY-1999;	99US-01322484P.	PR	22-JUL-1999;	99US-01450885P.
PR	05-MAY-1999;	99US-01322485P.	PR	22-JUL-1999;	99US-01450877P.
PR	06-MAY-1999;	99US-01322486P.	PR	22-JUL-1999;	99US-01450899P.
PR	06-MAY-1999;	99US-01322487P.	PR	22-JUL-1999;	99US-01451922P.
PR	07-MAY-1999;	99US-01322863P.	PR	23-JUL-1999;	99US-01451454P.
PR	11-MAY-1999;	99US-01342566P.	PR	23-JUL-1999;	99US-01452181P.
PR	14-MAY-1999;	99US-01342188P.	PR	23-JUL-1999;	99US-01452244P.
PR	14-MAY-1999;	99US-01342199P.	PR	26-JUL-1999;	99US-01452766P.
PR	14-MAY-1999;	99US-01343221P.	PR	27-JUL-1999;	99US-01459133P.
PR	14-MAY-1999;	99US-01343700P.	PR	27-JUL-1999;	99US-01459181P.
PR	18-MAY-1999;	99US-01343768P.	PR	27-JUL-1999;	99US-0145919P.
PR	19-MAY-1999;	99US-01344941P.	PR	28-JUL-1999;	99US-01459511P.
PR	20-MAY-1999;	99US-01355124P.	PR	02-AUG-1999;	99US-01463866P.
PR	21-MAY-1999;	99US-01353533P.	PR	02-AUG-1999;	99US-01463888P.
PR	24-MAY-1999;	99US-01356299P.	PR	02-AUG-1999;	99US-01463899P.
PR	25-MAY-1999;	99US-01360211P.	PR	03-AUG-1999;	99US-01470388P.
PR	25-MAY-1999;	99US-01363922P.	PR	04-AUG-1999;	99US-0147204P.
PR	28-MAY-1999;	99US-01367822P.	PR	04-AUG-1999;	99US-01473022P.
PR	01-JUN-1999;	99US-01372222P.	PR	05-AUG-1999;	99US-01471922P.
PR	03-JUN-1999;	99US-01375288P.	PR	05-AUG-1999;	99US-0147260P.
PR	04-JUN-1999;	99US-01375022P.	PR	06-AUG-1999;	99US-01473033P.
PR	07-JUN-1999;	99US-0137724P.	PR	06-AUG-1999;	99US-01474166P.
PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-01474933P.
PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138847P.	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	99US-0139119P.	PR	11-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	99US-01394522P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-01394533P.	PR	13-AUG-1999;	99US-0148565P.
PR	17-JUN-1999;	99US-01394532P.	PR	13-AUG-1999;	99US-0148684P.
PR	17-JUN-1999;	99US-01394922P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139454P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139455P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139456P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0150658P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-015066P.
PR	21-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151080P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	22-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0153663P.
PR	24-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99

Db 1390 COTCGTGCACGGATCAGTATATCGCCACATATCTAGC----- 1431
Qy 474 TyrArgAspArgSerPheProSerGlnArgAspGlnAspGluLeuSerAlaLeu 493
Db 1432 TTTTCTGACAAAGTTTC----- 1449
Qy 494 ValSerSerTyrLeuGlyProSerThrSerPheProHisArgSerArgSerProGlu 513
Db 1449 ----- 1449
Qy 514 TyrMetValProLeuProHisGlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeu 533
Db 1450 -----TACAGCCGCT 1461
Qy 534 AlaProAsnSerTyrSerProGlyHisGlyHisArgLeuHisArgGlnTyrSerProSer 553
Db 1462 GCGCCACAGAGG-----CATTCCTACCCCTTACGAGAGCGAGTAT----- 1500
Qy 554 LeuValHisGly-----GlnArgHisProLeuGlnTyrSerProProileHisGly 570
Db 1501 ---GTATATGGTGGCTGAGCCCATCATCCCAACATGATAAGC----- 1542
Qy 571 GlnGlnGlnLeuProTyrGlyGlnArgValTyrArgHisSerProSerGluArg 590
Db 1543 -----TCAGCTCCCTATGGCATG-----TCACCTGCCACACACACA 1578
Qy 591 TyrLeuGly 593
Db 1579 TACTATGGT 1587

RESULT 6

ADC03469

ID ADC03469 standard; DNA; 1602 BP.

XX AC ADC03469;

XX DT 18-DEC-2003 (first entry)

XX DE Maize flowering time-related DNA sequence #4.

XX KW flowering time; maize; flower architecture; plant biotechnology;

XX KW commercial plant farming; agriculture; flowering-related protein; ds.

XX OS Zea mays.

XX PN WO2003000904-A2.

XX PD 03-JAN-2003.

XX PF 24-JUN-2002; 2002WO-EP006968.

XX PR 22-JUN-2001; 2001US-0300112P.

XX PR 26-SEP-2001; 2001US-0325277P.

XX PR 30-NOV-2001; 2001US-0334984P.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Sessions A, Briggs S, Cooper B, Goff SA, Moughamer T;

XX PI Glazebrook J, Katagiri F, Krepis J, Provart N, Ricke D, Zhu T;

XX WPI; 2003-229340/22.

XX New isolated nucleic acid molecule encoding a polypeptide modulating

XX flower architecture and flowering time, useful in the area of plant

XX biotechnology, and commercial plant farming and agriculture.

XX Disclosure; SEQ ID NO 109; 323pp; English.

XX The invention comprises the amino acid and coding sequences of proteins

XX involved in the control of flowering time in rice. The DNA and protein

XX sequences of the invention are useful for modulating flower architecture

XX and flowering time, the DNA and protein sequences are useful in the area

XX

CC of plant biotechnology, commercial plant farming and agriculture. The
CC present maize DNA sequence shows homology to the rice sequences of the
CC invention.

SQ Sequence 1602 BP; 442 A; 336 C; 404 G; 420 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5 08e-17 Length: 1602
Score: 329.00 Matches: 142
Percent Similarity: 40.13% Conservative: 104
Best Local Similarity: 23.16% Mismatches: 195
Query Match: 10.60% Indels: 172
DB: 9 Gaps: 26

US-09-890-475-1 (1-609) x ADC03469 (1-1602)

Qy 36 ValGluThrGluSerThrSerMetPhePheThrLeuGlyGlnSerLysGlnProGlnPhe 55
Db 10 ATGGAGTCTGTGCGGCTCTTATGGACTCGACA-----AGCTCCAAGATACAACAGCTT 63
Qy 56 LeuLysSerIleAspGluLeuAlaAla----- 64
Db 64 CAGAGGCATTCGCTGAGCTTGAGCCAGAGCGCTTCTGCATGAACTTCAGTGGAG 123
Qy 65 -----PheSerValAlaValGluThrPheLysArgGlnPheAspLeu 79
Db 124 CAGCTTGAAGACCACCTTTCGTGCTTGAACAATCCCTCAAGAAAAAGTTTGCAGACTA 183
Qy 80 Gln----- 80
Db 184 AAAGACAGAGAGAGAGTTCAGGAGACTGTCGAAATCAGACAGATGCTGGAGCAG 243

Qy 81 -----LysHisIleGluSerIleGluSerIleGluAsnAlaIleAspSerLys 93
Db 244 CGGAGGCAGCTGTGTGGCCAGGAACAGCTCTTTGGAGAGGTTCAGAGAGAGAG 303

Qy 94 LeuGluSerAsnGlyValValLeuAlaAlaArgAsnAsnPheHisGlnProMetLeu 113
Db 304 GAAGCTGCATTAGCTATGATCTTTGCAAGTCCAGGCTTCTTACCTGTTCTTCATC 363

Qy 114 SerProProArg-----AsnAsnValSerValGluThrValThrValSerGln 130
Db 364 AACCCATGAACAAGCAGCTGGATAATCTTGGTGTCAAGTGGGCTAACCGGCTTCTGAG 423

Qy 131 ProSerGlnGluLe-----ValProGluThrSerAsnLysPro-----GluGlyGlyArg 147
Db 424 GAGAGTGTGCACCTGCAAGTTGATAATGCTGCAGTGCAGGCTGTTCTGAGCTTGTGCA 483

Qy 148 MetCysGluLeuMetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsnIleSerAsp 167
Db 484 CTCTGTGAGGAATGAATGAGACGGCTTCATAGTTTATATCCGACACAGGAGAAC 543

Qy 168 GlnAlaLysLeuMetGluGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLys 197
Db 544 TTGCGAGCCATCCGCGAGGAATTCCTAGTGCACTGAAGAAAACATCTCATCTTATGGC 603

Qy 188 PheValLeuAspCysIleGlyLysPheTyr-----LeuGlnGlyArg 201
Db 604 CTGTGTGTGATTCCTCTGGAGACTTTTATCTGGAGATAATCTTATATAGATGAGAAA 663

Qy 202 ArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuLeu 221
Db 664 -----AAGGATGGTGACCTTCTGGGTGTTAGAGGACATGTTTGTATGTTGATG 711

Qy 222 GluSerPheLeuLeuMetProAspArgGlyLysGlyVal----- 235
Db 712 GAGTCACCTT-----GGACACTGCACCTCTAATGATATACT 747

Qy 236 -----LysIleGluSerTrpIleLysAspGluAlaGluThrAla 248
Db 748 TGTCTTCTTTTGGAGAGCCACATGCTTACAACCTAATATCATAGAGCGCAAGAGGATT 807

Qy 249 AlaValAlaIleTrpArgLysArgLeu-----MetThrGluGlyGlyLeu 262

Db 808 GCATTTAAGTGAAGTCAAGTTAGATAATCTTGACATTCGATGCAATGGAACTGT 867
Qy 263 AlaalaalaGluLysMetAspAlaArgGlyLeuLeuLeuValaLalaCysPheGlyVal 282
Db 868 -----CTTGAAGCTCAGCGTTCTTCAATCTCGGCAACTTGTGATT 912
Qy 283 ProSerAsnPheArgSerThrAspLeuLeuAspLeuLeu-----ArgMetSer 298
Db 913 TCTGCTGAATTCATGAAGATGACTTGTGCAAAATGCTTCATATGTCAGTCGTCGCT 972
Qy 299 GlySerAsnGluLeuAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSer 318
Db 973 CAGACACTGAATCTTGTGATTCCTTGGGTTGTGGAG-----AAATGCCA 1020
Qy 319 GlyLeuValGluSerSerIleLysArgGlyMetHisIleGluAlaLeuGluMetValTyr 338
Db 1021 GGTGTCATTCGAATCTGCTGGAAGTGAAGACCTATTGATGCAATTAATTTGGCTTAT 1080
Qy 339 ThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMet 358
Db 1081 GTGTTTCAGCTTACTGAACATTTGAACCCAGTACACCTTCTTAAGCATATCTAAGGAT 1140
Qy 359 SerLysGluSerPheGluAlaLysArgLysAlaGlnSerProLeuAlaPheLysGlu 378
Db 1141 GTTAAAGAA---ANGTCACAGCCAGGAATGTCAAAACCTTCTCTGAGCAGACAGATGAG 1197
Qy 379 AlaAlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLys 398
Db 1198 ATGAATGAACGTGAGTATCTGCGCTGAAATCTGTCATTAAGTCATTAAGAGCAGACAAA 1257
Qy 399 LeuAspProAlaLysGluLeuProGlyTrpGlnLysGluGlnIleValSerLeuGlu 418
Db 1258 CTTGAG-----CAGCAGTACCTT-----GTGGATCCACITTCAG 1290
Qy 419 LysAspThrLeuGlnLeuAspLysGluMetGluLysAlaArgSerLeuSerLeuMet 438
Db 1291 AAAAGGGTTTTGCAGTAGAGAAAGCAAGCAGCAGAGGATGGCTGT----- 1341
Qy 439 GluGluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSer 458
Db 1342 ---GAGCAGCA-----AAGCCGAGTCCCAAGAGGCTCTGTGCCAAT 1380
Qy 459 ProMetGluMetProProValThrSerSerTyrSerProIleTyrArgAspArgSer 478
Db 1381 GGATCGGCCTTT-----GCACCCGCTGCTACAGGCTTGTGTCACAGAGC 1425
Qy 479 PheProSerGlnArgAspAspAspGlnAspGluLeuValSerSerTyrLeu 498
Db 1426 TTCTACTCA----- 1434
Qy 499 GlyProSerThrSerPheProHisArgSerArgSerProGluTyrMetValProLeu 518
Db 1435 -----GCAACCCAGAGAGGATCCATTCATCTCCGAGAG----- 1473
Qy 519 ProHisGlyGlyLeuGlyArgSerValTyrAlaTyrGlu-----HisLeuAlaPro----- 535
Db 1474 -----CAGTTGTGTATGCGCTGAGGTCATCTCTCCCATGATG 1515
Qy 536 -----AsnSerTyrSer-----ProGlyHisGlyHisArgLeuHisArgGlnTyrSer 551
Db 1516 ACCTCAGGCTTTACACAGTCAACCTGCCATGGACCT----- 1554
Qy 552 ProSerLeuValHisGlyGlnArgHisProLeuGlnTyr 564
Db 1555 -----TATTACGGTAATGGCTACCCAGTTCAGTAC 1584
RESULT 7
AAC49876
ID AAC49876 standard; DNA; 1964 BP.
XX
AC AAC49876;
XX

DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 62772.
XX Hybridisation assay; genetic mapping; gene expression control;
KW Protein identification; signal transduction pathway; metabolic pathway;
KW Promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0128785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.

PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	23-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152163P.
PR	24-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	28-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	29-JUN-1999;	99US-0141099P.	PR	15-SEP-1999;	99US-0154018P.
PR	30-JUN-1999;	99US-0141287P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0141842P.	PR	20-SEP-1999;	99US-0154779P.
PR	01-JUL-1999;	99US-0142154P.	PR	22-SEP-1999;	99US-0155139P.
PR	02-JUL-1999;	99US-0142055P.	PR	23-SEP-1999;	99US-0155486P.
PR	06-JUL-1999;	99US-0142390P.	PR	24-SEP-1999;	99US-0155659P.
PR	08-JUL-1999;	99US-0142803P.	PR	28-SEP-1999;	99US-0158458P.
PR	09-JUL-1999;	99US-0142920P.	PR	29-SEP-1999;	99US-0158596P.
PR	12-JUL-1999;	99US-0142977P.	PR	04-OCT-1999;	99US-0157117P.
PR	13-JUL-1999;	99US-0143542P.	PR	05-OCT-1999;	99US-0157753P.
PR	14-JUL-1999;	99US-0143624P.	PR	06-OCT-1999;	99US-0157865P.
PR	15-JUL-1999;	99US-0144005P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144085P.	PR	08-OCT-1999;	99US-0158232P.
PR	16-JUL-1999;	99US-0144086P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144323P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144333P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159638P.
PR	20-JUL-1999;	99US-0144884P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0144814P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160767P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161320P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161992P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
			Alignment Scores:		
			Pred. No.:	2,92e-14	Length: 1964
			Score:	296.50	Matches: 130
			Percent Similarity:	41.10%	Conservative: 101
			Best Local Similarity:	23.13%	Mismatches: 231
			Query Match:	9.55%	Indels: 101
			DB:	3	Gaps: 19
			US-09-890-475-1 (1-609) x AAC49876 (1-1964)		
Qy	59	IleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGln-----	75		
Db	128	TTAACTGAGTTAGCT-----CAGCCGAGTTTCGAGGAGTTTCAGAAACAGACGTCGTTG	181		
Qy	76	-----PheAspAspLeuGlnLysHisrIleGluSerIleGluAsn	88		
Db	182	ATGACTAGCTGTACTTCTTATCGCAAGAGCTTCGATCACTTCTCTTAGAGCAG	241		
Qy	89	AlaIleAspSerLysLeuGluSerAsnGlyValValLeuAlaAlaArgAsnAsnPhe	108		
Db	242	AACTGATGAAGAAATCTGAAGCGGTTGAAGCAGATGATTAACCTTAGATAACCACT	301		

XX 18-DEC-2003 (first entry)
 XX DE Rice flowering time-related gene #4.
 XX KW flowering time; rice; flower architecture; plant biotechnology;
 KW commercial plant farming; agriculture; flowering-related protein; gene;
 XX ds.
 XX OS *Oryza sativa*.
 XX PN WO2003000904-A2.
 XX PD 03-JAN-2003.
 XX PF 24-JUN-2002; 2002WO-BP006968.
 XX PR 22-JUN-2001; 2001US-0300112P.
 XX PR 26-SEP-2001; 2001US-0325277P.
 XX PR 30-NOV-2001; 2001US-0334984P.
 XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX PI Sessions A, Briggs S, Cooper B, Goff SA, Moughamer T;
 PI Glazebrook J, Katagiri F, Krepis J, Provart N, Ricke D, Zhu T;
 XX WPI; 2003-229340/22.
 XX DR P-FSDH; ADC03368.
 XX CC New isolated nucleic acid molecule encoding a polypeptide modulating
 PT flower architecture and flowering time, useful in the area of plant
 PT biotechnology, and commercial plant farming and agriculture.
 XX Claim 4; SEQ ID NO 7; 323pp; English.
 XX CC The invention comprises the amino acid and coding sequences of proteins
 CC involved in the control of flowering time in rice. The DNA and protein
 CC sequences of the invention are useful for modulating flower architecture
 CC and flowering time, the DNA and protein sequences are useful in the area
 CC of plant biotechnology, commercial plant farming and agriculture. The
 CC present DNA sequence encodes a rice flowering-related protein of the
 CC invention.
 XX SQ Sequence 1377 BP; 392 A; 275 C; 352 G; 358 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3 47e-14 Length: 1377
 Score: 293.00 Matches: 115
 Percent Similarity: 42.11% Conservative: 85
 Best Local Similarity: 24.21% Mismatches: 173
 Query Match: 9.44% Indels: 102
 DB: 9 Gaps: 19

US-09-890-475-1 (1-609) x ADC03367 (1-1377)

QY 36 ValGluThrGluSerThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPhe 55
 Db 10 ATGGAGTCCAGTCCACTATGAGTCCACA-----GGCTCCAAAGTCGACGAGCTT 63
 QY 56 LeuLysSerIleAspGluLeuAlaLa----- 64
 Db 64 CAGCGTGCAATTGCGAGGTCGAGAGCCGAGCGGTGTTCTTGAACCTAAAGTGAAG 123
 QY 65 -----PheSerValAlaValGluThrPheLysArgGlnPheAspAspLeu 79
 Db 124 CAACTGGAGGACCAATTTCCTGGGTGAGCAGTCGCTCAAGAAAGATTGATGATCTG 183
 QY 80 GlnLysHisIleGluSerIleGluAsnAlaIle-----AspSerLysLeuGluSer 96
 Db 184 AAAGGCAGAGAGAGAGGTTCCAGGAGACCGGTTGCAAGTCAAGATGCTGAGGCAA 243
 QY 97 AsnGlyValValLeuAlaLaArg----- 104

Db 244 CAAGAGGCTGTTGTCTGGCAAGAAAGTGAATCTTCTTGAATAAACTGCAGCAGAAAGG 303
 QY 105 -----AsnAsnAsnPheHisGlnProMetLeu 113
 Db 304 GATGCCGCTGGCTGTGATCTTTGGCAAGTCCAAAGTGAATCTTCTACGCTCTTATC 363
 QY 114 SerProProArgAsnAsnVal-----SerValGluThrThr 125
 Db 364 AACCAATCAGCAAAATCCGTTGAATAACAATGCTGTGTTTAAATGGAACATCGGTGATG 423
 QY 126 ValThrValSerGlnProSerGlnGluLeuVal-----ProGlu 138
 Db 424 CTTTCGTTAAGTGGCCAAACCTGCACTGGCATGGTGCATACCTGCAAGATGAAC 483
 QY 139 ThrSerAsnLysProGluGlyArg-----MetCysGluLeuMetCysSerLysGly 156
 Db 484 ACTGCTGTGAAGCCTCGTTCTCAACTTGTGTTCTCTGTGAAGAGATGAATGTTAACGG 543
 QY 157 LeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluPro 176
 Db 544 CTTCAATAGTTCAATACAGCAATCCAAAGACTTAACATCCATCCGTGAGGAATCCCC 603
 QY 177 SerAlaLeuLysLeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPhe 196
 Db 604 GTTGCACTTAGGGGAGCAACTGATCCATATGCTTGGTCTGCTTAGCTTCTCTAGAAGACTTC 663
 QY 197 Tyr-----LeuGlnGlyArgAlaPheThrLysGluSerProMet 210
 Db 664 TATTTGGAGACAATCTTATTTGGATGGCAAA-----AAGATGGAACCTT 711
 QY 211 SerSerAlaArgGlnValSerLeuLeuIleLeuGluSerPheLeu---LeuMetProAsp 229
 Db 712 CTGGGTGTACGGAGGAGCGCTCCCTCATGTTGATGAATCACTTCTCAGCTGCAAACTGAT 771
 QY 230 ArgGlyLysGlyLysValLys-----IleGluSerTrpIleLysAspGluAla 245
 Db 772 GCTACTACTGTTTCATATCGGAAGTCAAGTCAAGTCTTACTGCAAGATCAAGGCGTGCA 831
 QY 246 GluThrAlaAlaValAlaTrpArgLysArgLeu---MetThrGluGlyGlyLeuAlaLa 264
 Db 832 AAAAGAGTGTCTTGTGAGTGGAAATCCAAAGTGGATGAGCTTGACCTTGATGCTAGCAAT 891
 QY 265 AlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPheGlyValProSer 284
 Db 892 GGGAACTGCTCGAAGCACATCGCTTCTTCACTCTCGCTACCTTCTGATATTGCT 951
 QY 285 AsnPheArgSerThrAspLeuLeuAspLeuIle-----ArgMetSerGlySer 300
 Db 952 GAATTTGCACAGATGAGCTGTGTAAGTACTTCCCTCAGTCAGTCGTGCTGCAGACA 1011
 QY 301 AsnGluIleAlaGlyAlaLysArgSerGlnPheLeuValProMetValSerGlyIle 320
 Db 1012 CCTGAGCTTTGTAGAATCTTGGATTGTCCCAAGACATG---CCAGGTGCTCATTTGGAGT 1068
 QY 321 ValGluSerSerIleLysArgGlyMetHisIleGluAlaLeuMetValThrPhe 340
 Db 1069 CTG-----GTTGAATGGAGAACCTATTGATGCAATTAATTGCTTATGCAATT 1119
 QY 341 GlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLys 360
 Db 1120 GAGCTGACTAATCAGTTTGAACCGGTGAATTAATGAAAGCTTACTACAGGAGGTCAAG 1179
 QY 361 Glu-----SerPheGluArgAlaLysArgLysAlaGlnSerProLeuAlaPheLysGlu 378
 Db 1180 AGTGTGCCACATTTCAAGACTGGAAGATATCTCTGCAA-----GTACAGAAATGAA 1230
 QY 379 AlaAlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLys 398
 Db 1231 ATGAATGAACGAGAGCTGTCTGCTTGAAGAGCTGCTATTAAAGTGTATCGAGGAGCAAAA 1290
 QY 399 LeuAspProAlaLysGluLeuProGlyTrpGlnIleLysGluGlnIleValSerLeuGlu 418
 Db 1291 CTTGATGAG-----AAATATCCGATTGACCTGCTTCAG 1323

Oy 419 LysAspThrLeuGlnLeuAspLys-----GluMetGluGlu 430
Db 1324 AAAGAGTTATCCAGCTAGAGAAAGCTAAGGCAGACCAAGAGGAG 1368
RESULT 10
AAC38630
ID AAC38630 standard; DNA; 1865 BP.
AC AAC38630;
XX
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 21672.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.
XX 22-JUN-1999; 99US-0139899P.
XX 23-JUN-1999; 99US-0140353P.
XX 23-JUN-1999; 99US-0140354P.
XX 24-JUN-1999; 99US-0140695P.
XX 28-JUN-1999; 99US-0140823P.
XX 29-JUN-1999; 99US-0140991P.
XX 30-JUN-1999; 99US-0141287P.
XX 01-JUL-1999; 99US-0141842P.
XX 02-JUL-1999; 99US-0142154P.
XX 02-JUL-1999; 99US-0142055P.
XX 06-JUL-1999; 99US-0142390P.
XX 08-JUL-1999; 99US-0142803P.
XX 09-JUL-1999; 99US-0142920P.
XX 12-JUL-1999; 99US-0142977P.
XX 13-JUL-1999; 99US-0143542P.
XX 14-JUL-1999; 99US-0143624P.
XX 15-JUL-1999; 99US-0144005P.
XX 16-JUL-1999; 99US-0144085P.
XX 16-JUL-1999; 99US-0144086P.
XX 19-JUL-1999; 99US-0144325P.
XX 19-JUL-1999; 99US-0144331P.
XX 19-JUL-1999; 99US-0144332P.
XX 19-JUL-1999; 99US-0144333P.
XX 19-JUL-1999; 99US-0144334P.
XX 19-JUL-1999; 99US-0144335P.
XX 20-JUL-1999; 99US-0144352P.
XX 20-JUL-1999; 99US-0144632P.
XX 20-JUL-1999; 99US-0144884P.
XX 21-JUL-1999; 99US-0144814P.
XX 21-JUL-1999; 99US-0145086P.
XX 21-JUL-1999; 99US-0145088P.
XX 22-JUL-1999; 99US-0145085P.
XX 22-JUL-1999; 99US-0145087P.
XX 22-JUL-1999; 99US-0145089P.
XX 22-JUL-1999; 99US-0145192P.
XX 23-JUL-1999; 99US-0145145P.
XX 23-JUL-1999; 99US-0145218P.
XX 23-JUL-1999; 99US-0145224P.
XX 26-JUL-1999; 99US-0145276P.
XX 27-JUL-1999; 99US-0145913P.
XX 27-JUL-1999; 99US-0145918P.
XX 27-JUL-1999; 99US-0145919P.
XX 28-JUL-1999; 99US-0145951P.
XX 02-AUG-1999; 99US-0146386P.
XX 02-AUG-1999; 99US-0146388P.
XX 02-AUG-1999; 99US-0146389P.
XX 03-AUG-1999; 99US-0147038P.
XX 04-AUG-1999; 99US-0147204P.
XX 04-AUG-1999; 99US-0147302P.
XX 05-AUG-1999; 99US-0147192P.
XX 05-AUG-1999; 99US-0147260P.
XX 06-AUG-1999; 99US-0147302P.
XX 06-AUG-1999; 99US-0147416P.
XX 09-AUG-1999; 99US-0147493P.
XX 09-AUG-1999; 99US-0147935P.
XX 10-AUG-1999; 99US-0148171P.
XX 11-AUG-1999; 99US-0148319P.
XX 12-AUG-1999; 99US-0148341P.
XX 13-AUG-1999; 99US-0148565P.
XX 13-AUG-1999; 99US-0148684P.

Db 1079 AAAGCTTATCTGAGGATGCAAGAAAGCTACAGCTTTAATCACTGATGATTCACAAAT 1138
Qy 366 AlalysArglybAlaGlnSerProLeuAlaPheLysGluAlaAlaThrLysGlnLeuAla 385
Db 1139 TCTGGCGGATCTGGCATCTT-----GTGGCGCGCAAGAGCAATCA 1180
Qy 386 ValLeuSerValMetGlnCysMetGluThrHisLysLeuMetProAlaLysGluLeu 405
Db 1181 GCATTTAGAGCAGCTTAAATGATAGATAGAAATCAACTCGAG-----GAAGATTC 1234
Qy 406 ProGlyTrpGlnLysGluGlnLeuValSerLeuGluLysAspThrLeuGlnLeuAsp 425
Db 1235 CCACCCGAGATCTCAAGAGAGCTTGATCAGCTAGAGAGCAAAACCGAGAGAGA 1294
Qy 426 LysGluMetGluGluLysAlaArgSerLeuMetGluGluAlaAlaLeuAlaLys 445
Db 1295 AACACGCGGTGATCCGCCAAC----- 1318
Qy 446 ArgMetTyrAsnGlnGlnLeuLysArgProArgLeuSer---ProMetGluMetProPro 464
Db 1319 -----AAGAGAACCCGAGCAAGCTACAGCGGTCCAATGCCACCA 1357
Qy 465 ValThrSerSerSerTyrSerProLysTyrArgAspArgSerPheProSerGlnArgAsp 484
Db 1358 GCCAAAGCAGGCGGTATACCA----- 1378
Qy 485 AspAspGlnAspGluLeuSerAlaLeuValSerTyrLeuGlyProSerThrSerPhe 504
Db 1379 -----AAGCTTACGTCTCTCTTCCACCGCTCCACCAACATTC 1420
Qy 505 ProHisArgSerArgSerProGluTyrMetValProLeuProHisGlyGlyLeuGly 524
Db 1421 ATCAGATCCCAAGTCACTACCTCAATACGCTGTACCA----- 1459
Qy 525 ArgSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyr-----SerProGly 541
Db 1460 -----GATACACCATATCCCACTACCTACACCAACAGAGGCTTCCA 1507
Qy 542 HisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGly-----GlnArg 559
Db 1508 TAC-----CAATACTCACCGGCGGTTCATGGTCTTACCAAAACC 1549
Qy 560 HisProLeuGlnTyrSerProProLeuHisGlyGlnGlnLeuProTyrGlyLeu 578
Db 1550 TCTCGGTCAGTTAT-----CCGACAGCATATGGTACTACTCGATCCGGTGGCTGCCA 1606
Qy 579 -----GlnArgValTyrArgHisSerProSerGluArgTyrLeu 592
Db 1607 CCCCCTCCAGTTTACCATCTTCCCGCACCACCACTATATTCAGCATGCTTACTAC 1666
Qy 593 GlyLeuSerAsnGlnArgSerProArgSerAsnSerSerLeuAspPro 608
Db 1667 TGATTAGCAATGACAAAGGAAAAAAGAGATGGTAAAGTAAACCAACCA 1714

RESULT 11

ABZ13112

ID ABZ13112 standard; DNA; 1524 BP.

XX AC ABZ13112;

XX DT 21-JAN-2003 (first entry)

XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 917.

XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX OS Arabidopsis thaliana.

XX FN WO200216655-A2.

XX PD 28-FEB-2002.

XX PF 24-AUG-2001; 2001WO-US026685.

XX 24-AUG-2000; 2000US-022786SP.
PR 26-JAN-2001; 2001US-026464P.
PR 22-JUN-2001; 2001US-030011P.
XX (SCRI) SCRIPPS RES INST.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX Harper JF, Krops J, Wang X, Zhu T;
PI MPI; 2002-304127/34.
DR Identifying a stress condition to which a plant cell has been exposed and
XX producing plants with increased tolerance to these abiotic stresses.
XX Claim 144; SEQ ID NO 917; 577bp + Sequence Listing; English.
XX The invention relates to identifying a stress condition to which a plant
XX cell has been exposed, comprising: (a) contacting nucleic acid
XX representative of expressed polynucleotides in the plant cell with an
XX array or probes representative of the plant cell genome; and (b)
XX detecting a profile of expressed polynucleotides in the plant cell
XX characteristic of a stress response. The method is useful in the
XX production of transgenic plants, cells and seeds and in producing plants
XX with increased tolerance to abiotic stress. The present sequence is that
XX of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ1574) used
XX in methods of the invention. Note: The sequence data for this patent is
XX not represented in the printed specification but is based on sequence
XX information supplied to Derwent by the European Patent Office
SQ Sequence 1524 BP; 401 A; 342 C; 391 G; 390 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,25e-09	Length:	1524
Score:	234.50	Matches:	117
Percent Similarity:	34.38%	Conservative:	91
Best Local Similarity:	19.34%	Mismatches:	222
Query Match:	7.55%	Indels:	175
DB:	6	Gaps:	21

US-09-890-475-1 (1-609) x ABZ13112 (1-1524)

Qy 49 GlnSerLysGlnProGlnPheLeuLysSerIleAspGluLeuAlaPheSerValAla 68
Db 34 AAGTCCTCTCAGCGGAGTTCTTTGAG----- 60
Qy 69 ValGluThrPheLysArgGln-----PheAspAsp 78
Db 61 -----TTTCAGAAAGCAGCGCTCTCTAATCAGCAGCTGCAATCTTCTATGGAAGAG 111
Qy 79 LeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsnGly 98
Db 112 CTCTCTGAGCATTTCTCTCCATGGAGCAAAACCTGTAGTGAAGAAATCTGAGCTCTGAGA 171
Qy 99 ValValLeuAlaAlaArgAsnAsnPheHisGlnProMetLeuSerProProArgAsn 118
Db 172 CAGATGATTGAAACCCCTAGATAATCAACGACAGAGCTCGATTGAGTTACTCAAGCATCGT 231
Qy 119 AsnValSerValGluThrThrValThrValSer----- 129
Db 232 GAGGTGACTATAGACCAACGAGCTCGAGATCGCGGAGGAAAGTTGAGGAACGCGTTAGA 291
Qy 130 -----GlnProSerGlnGluIleValProGluThrSerAsnLysPro 143
Db 292 CGCGCGCTTGATTCGCTGGAGAAAGCTAGAGACTGTGGTGTAGTGAAGTACCGGAGAGTC 351
Qy 144 GluGlyGly-----ArgMetCysGluLeuMetCysSerLys 155
Db 352 GACGACGCGCAGCGGCTTTATTCGGCGCTCAAGTCGCTTTGTTGAAGTGGACGCAAGA 411
Qy 156 GlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIle 175
Db 412 CGATTTTGGGGTTCGTGATTGCGAGGAAGAGGATTCGAGAAATCTCCGGTCCACAGATT 471

Qy	515	MetValProLeuProHisGlyGlyLeuGlyArgSerValTYrAlaTYrGluHisLeuAla	534
Db	1306	GGTGTAACA-----GCATACACCATCCCA	1332
Qy	535	ProAenSerTYr-----SerProGlyHisGlyHisArgLeuHisArgGlnTYrSer	551
Db	1333	CTACCATCTACAGCAACAGGAGCCCTCCATAC-----CAATACTCA	1374
Qy	552	ProSerLeuValHisGly-----GlnArgHisProLeuGlnTYrSerProProIleHis	569
Db	1375	CCGAGAGCGGTTCAAGTTCCTTACCAAACTCTCCGTCAGTTAT---CCGACACATAT	1431
Qy	570	GlyGlnGlnGlnLeuProTYrGlyIle-----GlnArgValTYrArgHisSerProSer	587
Db	1432	GGTACTACTGCAGTCCGCTGCTCTCCACCCCTCCAGTTCACATCTCCACCGCAC	1491
Qy	588	GluGluArgTYrLeu	592
Db	1492	CACCAACCACCATATT	1506
RESULT	12		
ADC03447			
ID	ADC03447	standard; DNA; 947 BP.	
XX	AC	ADC03447;	
XX	AC		
XX	DT	18-DEC-2003 (first entry)	
XX	DT		
XX	XX	Wheat flowering time-related DNA sequence #7.	
XX	DE		
XX	DE	flowering time; wheat; flower architecture; plant biotechnology;	
XX	KW	commercial plant farming; agriculture; flowering-related protein; ds.	
XX	KW		
XX	OS	Triticum aestivum.	
XX	XX		
XX	PN	WC2003000904-A2.	
XX	PD	03-JAN-2003.	
XX	XX		
XX	PF	24-JUN-2002; 2002WO-EP006968.	
XX	XX		
XX	PR	22-JUN-2001; 2001US-0300112P.	
XX	PR	26-SEP-2001; 2001US-032527P.	
XX	PR	30-NOV-2001; 2001US-034984P.	
XX	XX		
XX	XX	(SYGN) SYNGENTA PARTICIPATIONS AG.	
XX	XX		
XX	PI	Sessions A, Briggs S, Cooper B, Goff SA, Moughamer T;	
XX	PI	Glazebrook J, Katagiri F, Kreps J, Provart N, Rickle D, Zhu T;	
XX	XX		
XX	XX	WFI; 2003-229340/22.	
XX	XX		
XX	PT	New isolated nucleic acid molecule encoding a polypeptide modulating	
XX	PT	flower architecture and flowering time, useful in the area of plant	
XX	PT	biotechnology, and commercial plant farming and agriculture.	
XX	XX		
XX	PS	Disclosure; SEQ ID NO 87; 323pp; English.	
XX	XX		
XX	CC	The invention comprises the amino acid and coding sequences of proteins	
XX	CC	involved in the control of flowering time in rice. The DNA and protein	
XX	CC	sequences of the invention are useful for modulating flower architecture	
XX	CC	and flowering time, the DNA and protein sequences are useful in the area	
XX	CC	of plant biotechnology, commercial plant farming and agriculture. The	
XX	CC	present wheat DNA sequence shows homology to the rice sequences of the	
XX	CC	invention.	
XX	XX		
XX	SQ	Sequence 947 BP; 248 A; 229 C; 231 G; 239 T; 0 U; 0 Other;	
Alignment Scores:			
Pred. No.:	8.4e-09	Length:	947
Score:	224.00	Matches:	82
Percent Similarity:	42.14%	Conservative:	52
Best Local Similarity:	25.79%	Mismatches:	100

176	ProSerAlaLeuLysLeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLys	195
472	CCTGTGGCGTTGGTAGATTGTGTGATCCGCGAAGTTAGTACTTGAAGCTGTATCGAG	531
196	PheTyrLeuGlnGlyArgArgAlaPheThrLysGluSerProMetSerSerAlaArgGln	215
532	CTTTTTCGGTGGATAAAGAGGT---CGTGGAGAGAAAGTGAGCAGATGATTTGGATGG	588
216	ValSerLeuLeuIleLeuGluSerPhe-----LeuLeuMetProAspArgLys	232
589	GCTTGTGTTGTGATCTCGAGAGTTTGATTCACGTTATGTTGATCCAGTCATGGGAAG	648
223	GlyLysValLysIleGluSerTrpIleLysAspGluAlaGluThrAlaAlaValAlaTrp	252
649	TCTCGCTACTGTAACTCCTAGCGCTTAAGAGAAAGCTAAAGAGATTGCTCGACAGCTGG	708
253	ArgLysArgLeuMetThrGluGlyLeuAlaAlaGluLysMetAspAlaArgGly	272
709	AAGGCTAGCTGGAAGAGAGAGAGGAGTGAAGATGTGAACACACCTGATGTTCATACG	768
273	LeuLeuLeuLeuValAlaCysPheGlyValProSerAsnPheArgSerThrAspLeuLeu	292
769	TTTTCGCAACATCTTGTGACTTTTCGGATGTGAAGAAG-----GATGATCTA	816
293	AspLeuIleArgMetSerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPhe	312
817	GCTCTTTATAGG-----AAACTGTGGTTGGTTTCAGCTTGGCGGTAAACAGATG	864
313	-----LeuValProMetValSerGlyIleValGluSerSer	324
865	CCAAAGCTTGCTGTTTCATGTTGGTGGCGACCAATGCCCTGATGATGTAAGAGTTG	924
325	IleLysArgGlyMetHisIleGluAlaLeuGluMetValTyrThrPheGlyMetGluAsp	344
925	ATPATCAGGGGCAACAGCTTGATCGGTTTCATTTTCACCTTTTGAAGTTGGTCTGTACAC	984
345	LysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGlu	364
985	TTGTTCCCTCCTGTTCTTGTCTCAAGCTTATCTGAGGGATGCAAGAAAGCACTACAGCT	1044
365	-----ArgAlaLysArgLysAlaGlnSerProLeu	374
1045	TTAATCACTGATGATCCAACTCTGCGCGATCTCGGGATCTGTTCTTCTGCATCTC	1104
375	AlaPheLysGluAlaAlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMet	394
1105	GCTTAC-----CAATGTATG	1119
395	GluThrHisLysLeuAspProAlaLysGluLeuProGlyTrpGlnIleLysGluGlnIle	414
1120	TGTTCTAGAGAGACCAAAACCCGAGAGAGAAACACGACC-----GTG	1161
415	ValSerLeuGluLysAspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSer	434
1162	ATTCCCGCCCAACAGAGAAACC-----	1182
435	LeuSerLeuMetGluGluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArg	454
1183	-----CGAGCAAGCTACAGCGGT-----	1200
455	ProArgLeuSerProMetGluMetProValThrSerSerSerTyrSerProIleTyr	474
1201	-----CCAATGCCACCGCCAAAGACGGCGGTATCACA-----	1233
475	ArgAspArgSerPheProSerGlnArgAspAspGlnAspGluIleSerAlaLeuVal	494
1234	-----AACGCTTACGTC	1245
495	SerSerTyrLeuGlyProSerThrSerPheProHisArgSerArgSerProGluTyr	514
1246	TCTCTTCCACCGGCTCCACCAACATTATCATAGATCCAAAGTCACTCACTCAATAC	1305

```
Query Match: 7.22% Indels: 84
DB: 9 Gaps: 14
US-09-890-475-1 (1-609) x ABC03447 (1-947)
Qy 268 MetAspAlaArgGlyLeuLeuValAlaCysPheGlyValProSerAsnPheArg 287
Db 10 TTGAAGCAGCATGATCTCTCACTGCTGCTACCTTGGTATTTCTGAAATTAAT 69
Qy 288 SerThrAspLeuLeuAspLeuLeu-----ArgMetSerGlySerAsnGluLeu 303
Db 70 GAAGATGAATCTTGCACAACTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 129
Qy 304 AlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyLeuValGluSer 323
Db 130 TGCCGACTGCTGGGTATATCAGAGATGCGAGGTGCTGATGAGTCTGGTGATAGT 189
Qy 324 SerLeuLysArgGlyMetHisLeuAlaLeuGluMetValTyrThrPheGlyMetGlu 343
Db 190 GCA-----AGACCAATGTGATGCAATTAACCTAGGCTACGGCTTGGCTCACT 237
Qy 344 AspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLysGluSerPhe 363
Db 238 GAACAGTTTGACCCAGTACATCTGTAAGCATATCTGAGGAGGTCACAGAAAG--GTG 294
Qy 364 GluArgAlaLysArgLysAlaGlnSerProLeuAlaPheLysGluAlaAlaThrLysGln 383
Db 295 TCACATGCCAAGATGGAATAATGCTCTCTGAGCAGCAGATGATGATGATGATGATG 354
Qy 384 LeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeuAspProAlaLys 403
Db 355 CTGCTGCACGAAAGTGTCTCAAGTGCATCAGGAGGACCAAACTGGAG-----GAG 408
Qy 404 GluLeuProGlyTyrGlnLeuLysGluGlnLeuValSerLeuGluLysAspThrLeuGln 423
Db 409 CAATATCCC-----GTGATCCACTTCAGAAAGGGTGATTCAG 447
Qy 424 LeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGluGluAlaLeu 443
Db 448 CTGAGAGGCGCAAGGAGCAGCAAGAGGAGGCTGTC-----GAAGCGCC-- 492
Qy 444 AlaLysArgMetTyrAsnGlnGlnLeuLysArgProArgLeuSerProMetGluMetPro 463
Db 493 -----AGCCACATCGAAGAGGCCACGT----- 516
Qy 464 ProValThrSerSerTyrSerProLeuTyrArgAspArgSerPheProSerGlnArg 483
Db 517 --GCCAAGCATCGCTATGCAACC-----CGTGTACAGCAGCTTCCCTGAGAAG-- 564
Qy 484 AspAspAspGlnAspGluLeuSerAlaLeuValSerSerTyrLeuGlyProSerThrSer 503
Db 565 -----AGCTTCTACCAAGGCAACACACAGAGAGG 591
Qy 504 PheProHisArgSerArgArgSerProGluTyrMetValProLeuProHisGlyGlyLeu 523
Db 592 CACCATACCTTACGAAAGGACGTTTGTGTAC----- 624
Qy 524 GlyArgSerValTyrAlaTyrGluHis-----LeuAlaProAsnSerTyr 538
Db 625 -----GCCGTGAGGCCACACACCTACGATGATGAACGGCGCTCCCTACCATC 678
Qy 539 SerProGlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGln 558
Db 679 TCACCCGCCAC-----ACGCGTACTACGGTAT 708
Qy 559 ArgHisProLeuGlnTyrSerProProLeuHisGlyGlnGlnLeuProTyr 576
Db 709 GGCTACCGCGTGCAGTAC-----CAGATACCTTAT 738
RESULT 13
ABL17653
ID ABL17653 standard; DNA; 7785 BP.
XX
```

AC ABL17653;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 4432.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX
XX Claim 1; SEQ ID NO 4432; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX sequences (AB101840-AB116175) and the encoded proteins (ABBS7737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 7785 BP; 2166 A; 2186 C; 2007 G; 1426 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.104 Length: 7785
Score: 152.00 Matches: 129
Percent Similarity: 37.78% Conservative: 106
Best Local Similarity: 20.74% Mismatches: 238
Query Match: 4.90% Indels: 149
DB: 4 Gaps: 31

US-09-890-475-1 (1-609) x ABL17653 (1-7785)
Qy 5 ProProThrValAlaAlaGlnProThr---ThrThrAlaAsnProLeuGlnArgHis 23
Db 3487 CTCCTGTTGGTGGGCAAGTCCAGCCCGCCGCAAGCCATCTTTC----- 3534
Qy 24 GlnSerGluGlnArgArgGluLeuProLysLeuValGluThrGluSerThr----- 41
Db 3535 -----GACCAACCCCGCACCAGCAGAGCCCGTCAATCCACCTATGCCACCCAGT 3588
Qy 42 -----SerMetAspIleThrIleGlyGlnSerLys 51
Db 3589 CCTATCAGATGGAGCAGCAGCACCCTGAAGAACCTGGAGATCTATTGGCAACATGGA 3648
Qy 52 GlnProGlnPheLeuLysSerIleAspGluLeuAlaAlaPheSerValAlaValGluThr 71
Db 3649 CAGAACCGCCCAAGCAGAACCCAGGAGGTCTACGTCTCCATCGAGACCATGCTGCACA 3708
Qy 72 PheLysArgGlnPheAspAspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAsp 91

XX New polypeptides that phosphorylate kinase, used to screen for modulators
 PT for treating e.g. cancer or inflammation.
 XX
 XX Claim 12; Fig 1; 95pp; English.
 XX
 CC The present sequence encodes rat TAO1 protein kinase, which is capable of
 CC phosphorylating MEK3 (a MAP/ERK kinase). TAO kinases, and related
 CC polypeptides, are used to screen for modulators of stress-responsive
 CC mitogen activated protein (MAP) kinase pathways. These modulators are
 CC potentially useful for treating or preventing: (1) inflammation,
 CC autoimmune disease, cancer and degeneration (inhibitors of
 CC phosphorylation); or (2) insulin-resistant diabetes, metabolic disorders
 CC and neurodegeneration (enhancers of phosphorylation). TAO kinases are
 CC also used to raise specific antibodies, useful therapeutically as
 CC modulators and as immunoassay reagents for detecting TAO kinases. TAO
 CC kinase polynucleotides can be used: (a) for recombinant expression of TAO
 CC kinases; and (b) in the form of fragments, for detecting TAO kinase
 CC polynucleotides in standard hybridisation and amplification tests. TAO
 CC kinases are highly specific for MEK3
 XX
 SQ Sequence 3312 BP; 1106 A; 674 C; 780 G; 752 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0.0391 Length: 3312
 Score: 151.00 Matches: 150
 Percent Similarity: 34.75% Conservative: 95
 Best Local Similarity: 21.28% Mismatches: 237
 Query Match: 4.86% Indels: 223
 DB: 2 Gaps: 32

US-09-890-475-1 (1-609) x AAZ32435 (1-3312)

Qy	3	AsnTyr-----	ProProThrValAlaAlaGlnProThrThrAlaAsn	17
Db	1333	AATTACCAAGAAGAGGAGATCTCTAGAACAGAGCATCAGCTCCACAGTCTCCA-----	1386	
Qy	18	ProLeuLeuGlnArgHisGlnSerGluGlnArgArgGluLeuProLysIleValGlu	37	
Db	1387	CCTCAAGTGTCGTGTCAAAATCATCATATGCTGTAATAGAGAACACTTGGCAACTATACGA	1446	
Qy	38	ThrGluSerThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLys	57	
Db	1447	ACAGCATCATCTG-----	1485	
Qy	58	SerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAsp	77	
Db	1486	CAGGACTCTGAACCTTAGAGAACAGATGTCGTATTATAGCGGATGAGCGGCACAGCAT---	1542	
Qy	78	AspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsn	97	
Db	1543	-----CAGAAAGCAGCTGTAGCTCTGTGAAATAATAACTGAAGGCAGAAATGGACGAACAT	1596	
Qy	98	GlyVal-----	113	
Db	1597	CGGCTCAGATTAGACAAAGATCTTGAACCTCAGCGCAACATTTTCGTGCAGAAATG---	1653	
Qy	114	SerProProArgAsnAsnValSerValGluThrThrValThrValSerGlnProSerGln	133	
Db	1654	-----GAGAAACTTATTAAAGAACCAACCAAGCTTCTATG	1686	
Qy	134	GluIleValProGluThrSerAsnLysProGluGlyGlyArgMetCysGluLeuMetCys	153	
Db	1687	GAATAAGAGGCTTAAGTGATGCCCAACGAGAG-----	1719	
Qy	154	SerLysGlyLeuArgLysTyvIleTyrAlaAlaSerAsnIleSerAspGlnAlaLysLeuMetGlu	173	
Db	1720	---AAAAAATTCACACACATTT-----CAGGCTCAACAGAGAA	1758	
Qy	174	GluIleProSerAlaLeuLys-LeuAlaLysGluProAlaLysPheValLeuAspCysIle	193	
Db	1759	GAATGATAGCTTTTGGATCTCAAAAAAGAGAAATATAAATCTCGAAAAGAGCACTT	1818	

```
Db 2803 CTCTCCCTCAGGCATTGAGCCACAGCTACCCAGGA----- 2838
Qy 490 IleserAlaLeuValSerSerTyrLeuGlyProSerThrSerPheProHisArgSerArg 509
Db 2839 -----CCTTCTAGCTGGTCTCACAAAT----- 2859
Qy 510 ArgSerProGluTyrMetValProLeuProHisGlyLeuGlyArgSerValTyrAla 529
Db 2860 -----CCTACTGGGGTTCAGGA----- 2877
Qy 530 TyrGluHisLeuAlaProAsnSerTyrSerProGlyHisGlyHisArgLeuHisArgGln 549
Db 2878 -----CCTCACTGGGGTCACTCCCATG----- 2898
Qy 550 TyrSerProSerLeuValHisGlyGlnArgHisProLeuGlnTyrSerPro----- 566
Db 2899 -----GCTGGCACACCAAGCTTGGGGTCACTCCGATGCAAGCGGACCCCAACCATGG 2952
Qy 567 -----ProIleHisGlyGlnGlnLeuProTyrGlyTyrGlyGlnArgValTyrArgHis 584
Db 2953 GGTCACTCCCTCAGGGCCAAATCAAGGGGTACCTCGAGTAGCAGTATAGGATCCGCAAT 3012
Qy 585 SerProSerGluGluTyrLeu-----GlyLeuSerAsnGlnArgSerProArg 601
Db 3013 AGCCCCAGGCTCTCAGGCGGACAGCTTCTGGGGGACGAGCGGAACAGGGCATGAGCAGA 3072
Qy 602 SerAsnSer 604
Db 3073 AGCAGGAGT 3081

RESULT 15
ABT40852
ID ABT40852 standard; DNA; 3312 BP.
XX AC ABT40852;
XX DT 26-JUN-2003 (first entry)
XX DE Toxicity modelling related rat gene SEQ ID No 554.
XX KW Toxic effect; gene expression profile; renal toxicity; toxicity marker;
XX KW database; drug screening; toxicity assay; rat; ds.
XX OS Rattus norvegicus.
XX PN WO200295000-A2.
XX PD 28-NOV-2002.
XX PF 22-MAY-2002; 2002WO-US016173.
XX PR 22-MAY-2001; 2001US-0292335P.
XX PR 13-JUN-2001; 2001US-0297523P.
XX PR 19-JUN-2001; 2001US-0298252P.
XX PR 10-JUL-2001; 2001US-0303807P.
XX PR 10-JUL-2001; 2001US-0303808P.
XX PR 28-AUG-2001; 2001US-0315047P.
XX PR 27-SEP-2001; 2001US-0324928P.
XX PR 22-OCT-2001; 2001US-0330462P.
XX PR 01-NOV-2001; 2001US-0330867P.
XX PR 21-NOV-2001; 2001US-0331805P.
XX PR 06-DEC-2001; 2001US-0336144P.
XX PR 19-DEC-2001; 2001US-0340873P.
XX PR 21-FEB-2002; 2002US-0357842P.
XX PR 21-FEB-2002; 2002US-0357843P.
XX PR 21-FEB-2002; 2002US-0357844P.
XX PR 15-MAR-2002; 2002US-0357844P.
XX PR 08-APR-2002; 2002US-0370144P.
XX PR 08-APR-2002; 2002US-0370206P.
XX PR 08-APR-2002; 2002US-0370247P.
XX PR 17-APR-2002; 2002US-0372794P.
```

```
PR 21-APR-2002; 2002US-0371679P.
XX (GENE-) GENE LOGIC INC.
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
XX WPI; 2003-148464/14.
XX Predicting at least one toxic effect of a compound, useful for toxicity
PT modeling, comprises preparing a gene expression profile of a tissue or
PT cell sample exposed to the compound, and comparing the gene expression
PT profile to a database.
XX Example 4; Page; 446pp; English.
XX The invention relates to a novel method of predicting at least one toxic
CC effect of a compound. The method comprises a gene expression profile of a
CC tissue or cell sample exposed to the compound, and comparing the gene
CC expression profile to a database comprising at least part of the data or
CC information given in the specification. The methods are useful for
CC predicting at least one toxic effect of a compound, predicting the renal
CC progression of a toxic effect of a compound, predicting the renal
CC toxicity of a compound, or identifying toxicity markers in tissues or
CC cells exposed to known renal toxin. The genes are useful as toxicity
CC markers in drug screening and toxicity assays, in monitoring disease or
CC physiological states, or disease progression. This polynucleotide
CC represents a rat DNA sequence relating to the toxic effect database
CC described in the specification. NOTE: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the World Intellectual Property
CC Organization
XX SQ Sequence 3312 BP; 1106 A; 674 C; 780 G; 752 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.0391 Length: 3312
Score: 151.00 Matches: 150
Percent Similarity: 34.75% Conservative: 95
Best Local Similarity: 21.28% Mismatches: 237
Query Match: 4.86% Indels: 223
DB: 7 Gaps: 32
```

US-09-890-475-1 (1-609) x ABT40852 (1-3312)

```
Qy 3 AsnTyr-----ProProThrValAlaAlaGlnProThrThrThrAlaAsn 17
Db 1333 AATTACCAAGAAGAGGAGATCTAGACACAGAGCATCAGTCCACAGTCTCCA----- 1386
Qy 18 ProLeuLeuGlnArgHisGlnSerGluGlnArgArgGluLeuProLysIleValGlu 37
Db 1387 CCTCAAGTGTCTGTCACAAATCAATATCTGTAATAGAGAACACTTTGCCACTATACGA 1446
Qy 38 ThrGluSerThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLys 57
Db 1447 ACAGCATCACTG-----GTTACAAGACAGATGCAAGACATGAG 1485
Qy 58 SerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAsp 77
Db 1486 CAGGACTCTGAATCTAGAACACAGATGCTGTTATAGCGGATGAGGCGAGCAT---- 1542
Qy 78 AspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsn 97
Db 1543 -----CAGAGCAGCTGATGATCTCGGAAATAAATACTGAAGCGAGAAATGCGCAACAT 1596
Qy 98 GlyVal-----ValleuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeu 113
Db 1597 CGGCTCAGATTAGACAAAGATCTTGAACACTCGCCGCAACAACTTCGCTCGCAAAATG--- 1653
Qy 114 SerProArgAsnAsnValSerValGluThrThrValThrValSerGlnProSerGln 133
Db 1654 -----GAGAAACTTATTAGAAACACCAAGCTTCTATG 1686
Qy 134 GluIleValProGluThrSerAsnLysProGluGlyGlyArgMetCysGluLeuMetCys 153
```



```
Db 1687 GAAAAAGAGCTAAAGTGTGATGGCAACAGAGAG----- 1719
Qy 154 SerLysGlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGlu 173
Db 1720 ---AABAAATTCACACACATTT-----CAGGCTCAACACAGAGAA 1758
Qy 174 GluIleProSerAlaLeuLys-LeuAlaLysGluProAlaLysPheValLeuAspCysI1 193
Db 1759 GAACGTGAATAGCTTTTGGAGTCTCAAAAAGAGAGATATAAATCTCGAAAGAGCAGCTT 1818
Qy 193 eGlyPhePheTyrLeuGlnGlyArgAlaPheThrLysGluSer-----PrometSe 211
Db 1819 AAGAGAGAGCTGAATGAACACAGACACACCTTAATAAAGAGAGAGGATGGCTTCA 1878
Qy 211 rSerAlaArgGlnValSerLeuLeuIleLeuGluSerPheLeuLeuMetProAspArgG1 231
Db 1879 AAGCAGAAGAGAGATATTCAACATTTTCAGGCACAGAGAGAGCTATCTTCTTCGACGT 1938
Qy 231 YLysGlyLysVal-----LysIleGluSerTyrPheLysAspGluAlaGluTh 247
Db 1939 CAAGGAGAGTATCTAGAGTAGATGTGTCGCTTCA-----AAGAAGATGTACTT 1992
Qy 247 rAlaAlaValAlaTyrArgLysArgLeuMetThrGlu-GlyGlyLeuAlaAlaGluL 267
Db 1993 GGTCCGCATAACTTGAACAGAGACCTTG-----TCAGGAGAGAGTT-----AAACAAA 2040
Qy 267 yMetAspAlaArgGlyLeuLeuLeu-----LeuValAlaLac 279
Db 2041 AGGCAGAGCTCAAGAGGACTTAGAACATGCATGTACTTGGACAGCAGATGAATCCTGCAA 2100
Qy 279 yspheGlyVal---ProSerAsnPheArgSerThrAspLeuLeu-AspLeuIleArgMet 297
Db 2101 GAACGTGGAGTTTCCGCCACCTCAACACTATTTCAGAGAGATGGCTGTGAGTTGATCAGACTG 2160
Qy 298 SerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValPrometVal 317
Db 2161 CAACATCAAACTGAGCTTACTTAACAGCTGGAATACATAAAGAGAGAGG----- 2208
Qy 318 SerGlyIleValGluSerSerIleLysArgGlyMethIleGluAlaLeuGluMetVal 337
Db 2209 -----GAACGGGAAGTAAAGCGG---AAACATGTC----- 2235
Qy 338 TyrThrPheGlyMetGluAspLysPheSerAlaLeuValLeuThrSerPheLeuLys 357
Db 2236 -----ATGGAAGTTTCGACAGCAGCTTAAGAGTTTGAAGTCTAAAGAACTCAA 2283
Qy 358 MetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeuAla----- 375
Db 2284 ATAAAAAGCAGTTTCAGGATACCTGCNAATTCAAACACAGACAGTACAAAGCATTAAGG 2343
Qy 376 -----PheLys 377
Db 2344 AATCACCTACTGGAGACTTACACCAAGAGTGAGCACAAGAGTGTCTGAAAAGAGACTCAAG 2403
Qy 378 GluAlaAlaThrLysGlnLeuAlaValLeuSer-----SerValMetGln 392
Db 2404 GAGGAACAGACTCGGAAGTTAGCCATCTTGGCTGAGCAGTATGATCATGATTAATGAA 2463
Qy 393 CysMetGluThrHis-----LysLeuAspProAlaLysGlu----- 404
Db 2464 ATGCTCTCCACACAAGCTCTCGGTTGGATGAAGCAGGAGCAGAGTGCAGGTTTGG 2523
Qy 405 -----LeuProGlyTyrGlnIleLys----- 411
Db 2524 AAGATGCAGCTACAGCAGGAATCGAGCTGTGTAATGCATATCAGACAAATCAAGATG 2583
Qy 412 -----GluGlnIleValSerLeu--- 417
Db 2584 CAGGCTGAGGCCCAACATGATCGAGAGCTTCGAGAGCTGGACAAAGGGTCTCCCTTCGG 2643
Qy 418 -----GluLysAspThrLeuGlnLeuAspLysGluMetGlu 429
```

```
Db 2644 AGACACACTCTTAGACACAGAGATTGAAGAGAGATGTTGGCTTTGCAGATGAAGGCACA 2703
Qy 430 GluLysAlaArgSerLeuSerLeuMetGluGluAlaAlaLeuAlaLysArgMetTyrAsn 449
Db 2704 GAACGAATACCT-----AGCCTGCTCGAGCCGACAGAGAAATTTGAAGCTTTTGAC 2757
Qy 450 GlnGlnIleLysArgProArgLeuSerProMetGluMetProProValThrSerSerSer 469
Db 2758 TCTGAAGCATCAGATTAGGTTTAACTAGTGCCTT-----TCTAAT 2802
Qy 470 TyrSerProIleTyrArgAspArgSerPheProSerGlnArgAspAspGlnAspGlu 489
Db 2803 CTCTCCCTCAGGCATTCAGCCACAGCTACCCAGGA----- 2838
Qy 490 IleSerAlaLeuValSerSerTyrLeuGlyProSerThrSerPheProHisArgSerArg 509
Db 2839 -----GCTTCTAGCTGGTCTCACAAT----- 2859
Qy 510 ArgSerProGluTyrMetValProLeuProHisGlyLeuGlyArgSerValTyrAla 529
Db 2860 -----CCTACTGGGGGTTTCAGGA----- 2877
Qy 530 TyrGluHisLeuAlaProAsnSerTyrSerProGlyHisGlyHisArgLeuHisArgGln 549
Db 2878 -----CCTCACTGGGGTCAATCCCATG----- 2898
Qy 550 TyrSerProSerLeuValHisGlyGlnArgHisProLeuGlnTyrSerPro----- 566
Db 2899 -----GGTGGCACACCAACAGCTTGGGGTTCATCCGATGCAAGGGGACCCCAACCATGG 2952
Qy 567 -----ProIleHisGlyGlnGlnLeuProTyrGlyIleGlnArgValTyrArgHis 584
Db 2953 GGTCAACCCCTCAGGGCCCAATGCAGGGGTACCTCGAGGTAGCAGTATAGGAGTCCGCAAT 3012
Qy 585 SerProSerGluGluArgTyrLeu-----GlyLeuSerAsnGlnArgSerProArg 601
Db 3013 AGCCCCCAGGCTCTGAGGCGGACACCTTCTGGGGACGACGGAACAGGCGATGAGCACA 3072
Qy 602 SerAsnSer 604
Db 3073 AGCAGCAGT 3081
```

Search completed: February 28, 2004, 03:15:01
Job time : 618 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 28, 2004, 00:19:58 ; Search time 131 Seconds
(without alignments)
2579.886 Million cell updates/sec

Title: US-09-890-475-1

Perfect score: 3104

Sequence: 1 MSNPPVVAQPTTANPLL.....RYLGLSNQSPRNSLDPK 609

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp
-Q=/cgn2_1/USPRO_spool_p/US9890475/runat_24022004_135312_5317/app_query.fasta_1.775
-DB=Issued Patents_NA -QFWT=frascap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blcsum62 -TRANS=human4.0.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US9890475 @CGN 1 1 103 @runat_24022004_135312_5317 -NCPU=6 -ICPU=3
-NO MWAP -LARGESQUERY -NEG SCORES=0 -WALT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/FACTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	151	4.9	3312	3	US-09-060-410-1
2	151	4.9	3312	4	US-09-723-458-1
3	148.5	4.8	3824	4	US-09-688-188B-20
4	148.5	4.8	3824	4	US-09-291-417D-20
5	137.5	4.4	6558	4	US-09-491-356C-7
6	133	4.3	17612	3	US-08-911-853-29
7	133	4.3	17612	3	US-09-479-409-29
8	133	4.3	17612	4	US-09-479-453-29
9	131	4.2	9775	3	US-08-977-171-1
10	130.5	4.2	7393	4	US-09-620-312D-372
11	129	4.2	3324	4	US-09-620-312D-1020
12	127.5	4.1	3605	3	US-09-098-901-1

13	127.5	4.1	10432	4	US-09-919-172-97	Sequence 97, Appl
14	127.5	4.1	10432	4	US-09-978-594-21	Sequence 21, Appl
15	127	4.1	6306	1	US-08-466-390-3	Sequence 3, Appl
16	127	4.1	6306	1	US-08-470-950-3	Sequence 3, Appl
17	127	4.1	6306	1	US-08-467-781-3	Sequence 3, Appl
18	127	4.1	6306	1	US-08-195-487-3	Sequence 3, Appl
19	127	4.1	6306	2	US-08-483-924-3	Sequence 3, Appl
20	127	4.1	6306	5	PCT-US93-06160-3	Sequence 3, Appl
21	126.5	4.1	9334	3	US-08-977-171-2	Sequence 2, Appl
22	122.5	3.9	7100	4	US-09-308-375-1	Sequence 1, Appl
23	121.5	3.9	6436	4	US-09-600-099-1	Sequence 1, Appl
24	120	3.9	11220	3	US-09-105-537-32	Sequence 32, Appl
25	120	3.9	36778	3	US-09-105-537-5	Sequence 5, Appl
26	119.5	3.8	5053	3	US-08-685-576-2	Sequence 2, Appl
27	118	3.8	38506	4	US-09-320-878-19	Sequence 19, Appl
28	118	3.8	38506	4	US-09-141-908-1	Sequence 1, Appl
29	118	3.8	38506	4	US-09-657-440-19	Sequence 19, Appl
30	117.5	3.8	2487	1	US-08-261-304-1	Sequence 1, Appl
31	116.5	3.8	4403765	3	US-09-103-840A-2	Sequence 2, Appl
32	116.5	3.8	4411529	3	US-09-103-840A-1	Sequence 1, Appl
33	116	3.7	1512	3	US-08-911-853-6	Sequence 6, Appl
34	116	3.7	1512	3	US-09-479-403-6	Sequence 6, Appl
35	116	3.7	1512	4	US-09-479-453-6	Sequence 6, Appl
36	116	3.7	2680	2	US-08-533-306A-5	Sequence 5, Appl
37	116	3.7	2680	2	US-08-742-923A-5	Sequence 5, Appl
38	116	3.7	2887	2	US-08-533-306A-3	Sequence 3, Appl
39	116	3.7	2887	2	US-08-743-923A-3	Sequence 3, Appl
40	116	3.7	7453	4	US-09-620-312D-248	Sequence 248, App
41	116	3.7	7501	4	US-09-620-312D-249	Sequence 249, App
42	116	3.7	10136	1	US-08-353-700-2	Sequence 2, Appl
43	116	3.7	10136	5	PCT-US95-16216-2	Sequence 2, Appl
44	115.5	3.7	2859	4	US-09-252-991A-944	Sequence 944, App
45	115.5	3.7	9626	4	US-09-150-867-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-060-410-1
; Sequence 1, Application US/09060410
; Patent No. 6165461
; GENERAL INFORMATION:
; APPLICANT: Cobb, Melanie
; APPLICANT: Hutchinson, Michele
; APPLICANT: Chen, Zhu
; APPLICANT: Berman, Kevin
; TITLE OF INVENTION: TAO PROTEIN KINASES AND METHODS OF USE
; TITLE OF INVENTION: THEREFOR
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,410
; FILING DATE: 14-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 860098.421
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3312 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 121..3123
 US-09-060-410-1

Alignment Scores:

Pred. No.: 7,19e-05 Length: 3312
 Score: 151.00 Matches: 150
 Percent Similarity: 34.75% Conservative: 95
 Best Local Similarity: 21.28% Mismatches: 237
 Query Match: 4.86% Indels: 223
 DB: 3 Gaps: 32

US-09-890-475-1 (1-609) x US-09-060-410-1 (1-3312)

QY 3 AsnTyr-----ProThrValAlaAlaGlnProThrThrAlaAsn 17
 DB 1333 AATTACCAAGAAAGGAGATCTTAGAACAGAGCATCAGCTCCACAGCTCTCCA-----1386
 QY 18 ProLeuLeuGlnArgHisGlnSerGluGlnArgArgGluLeuProProLysIleValGlu 37
 DB 1387 CCTCAAGTCTCTGTCACAAATCACATTATCGTAATAGAACACATTTCGCAACTATACGA 1446
 QY 38 ThrGluSerThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLys 57
 DB 1447 ACAGCATCATCTG-----GTTACAAGACAGATCAAGAACATCAG 1485
 QY 58 SerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAsp 77
 DB 1486 CAGGACTCTGAATCTAGAGAACAGATGCTGCTTTAAGCGGATGAGCGCAGCAT---1542
 QY 78 AspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsn 97
 DB 1543 -----CAGAAGCAGCTGATGACTCTGGAAATTAACCTGAAGCGCAGAAATGGACGAACAT 1596
 QY 98 GlyVal-----ValLeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeu 113
 DB 1597 CGGCTCAGATTAGACAAAGATCTTGAACCTCAGCGCAACAATTCGCTGCAGAAATG---1653
 QY 114 SerProArgAsnAsnValSerValGluThrThrValThrValSerGlnProSerGln 133
 DB 1654 -----GAGAACTTATTAGAACACCAACAGCTTCTATG 1686
 QY 134 GluIleValProGluThrSerAsnLysProGluGlyGlyArgMetCysGluLeuMetCys 153
 DB 1687 GAAAAGAGGCTAAAGTATGTCGCCAACAGAGGAG-----1719
 QY 154 SerLysGlyLeuArgLysThrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGlu 173
 DB 1720 -----AAAAATTCCAACACACAT-----CAGCTCAACAGAGAA 1758
 QY 174 GluIleProSerAlaLeuLys-LeuAlaLysGluProAlaLysPheValLeuAspCysIle 193
 DB 1759 GAACCTGAATAGCTTTTGGAGTCTCAAAAAGAGAATATAAACTTCGAAAAGAGCAGCTT 1818
 QY 193 eGlyLysPheTyrLeuGlnGlyArgAlaPheThrLysGluSer-----ProMetSe 211
 DB 1819 AAGGAGGAGCTGAATGAAACACAGACACACCTTAAAAAAGAAAGAGGAATGGCTTTCA 1878
 QY 211 rSerAlaArgGlnValSerLeuLeuIleLeuGluSerPheLeuLeuMetProAspArgIle 231
 DB 1879 AAGCAGAAGGAGATATTCAACATTTTCAGGCGAGAGAGAGAGCTAATCTTCTCGACGT 1938
 QY 231 yLysGlyLysVal-----LysIleGluSerTrpIleLysAspGluAlaGluTh 247
 DB 1939 CAAAGCAGTATCTAGAGCTAGAATGTCGTCCTTCA-----AAAGAGATGTTACTT 1992
 QY 247 rAlaAlaValAlaTrpArgLysArgLeuMetThrGlu-GlyGlyLeuAlaAlaGluL 267

DB 1993 GGTGGCATAAAGTGGAAACAGGACCTTG-----TCAGGAGAGGTT-----AAACAAA 2040
 QY 267 yMetAspAlaArgGlyLeuLeu-----LeuValAlaC 279
 DB 2041 AGCAGACTCAGAGAGCTTAGAACATGCAATGTACTCGCAGACAGCATGAATCCATGCAA 2100
 QY 279 yPheGlyVal---ProSerAsnPheArgSerThrAspLeuLeu-AspLeuIleArgMet 297
 DB 2101 GAACCTGGAGTTTCGCCACCTCAACACTATTTCAGAAAGATCGCTGTGAGTTGATCAGACTG 2160
 QY 298 SerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetVal 317
 DB 2161 CAACATCAAACTAGCTTACTACACAGCTGGAAATCAATAAGAGAAGG-----2208
 QY 318 SerGlyIleValGluSerSerIleLysArgGlyMetHisIleGluAlaLeuGluMetVal 337
 DB 2209 -----GAACGGGAACCTAAGACGG---AAACATGTC-----2235
 QY 338 TyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLys 357
 DB 2236 -----ATGAAGTTTCGACAGCAGCTTAAGAGTTTGAAGTCTAAAGAACTCCAA 2283
 QY 358 MetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeuAla-----375
 DB 2284 ATAAAGAGCAGTTTCAGGATACCTGCAAAATTCAAACCCAGACAGTACAAAGCATTAAGG 2343
 QY 376 -----PheLys 377
 DB 2344 AATCACCCTACTGGAGACTACACAAAGACTGAGCACAAAGCTGTTCTGAAAAGACTCAAG 2403
 QY 378 GluAlaAlaThrLysGlnLeuAlaValLeuSer-----SerValMetGln 392
 DB 2404 GAGGAACAGCTCGGAAGTTAGCCATCTTGGCTGACGACATGATCATAGCATTATAGAA 2463
 QY 393 CysMetGluThrHis-----LysLeuAspProAlaLysGlu-----404
 DB 2464 ATGCTCTCCACACAACAGCTGCTGCTTTGGATGAAGCACAGCAAGCAGATGCCAGGTTTG 2523
 QY 405 -----LeuProGlyTyrGlnIleLys-----411
 DB 2524 AGATGACGTACAGCAGGAACTGGAGCTGTTGATGCATATCAGAGCAAAATCAAGATG 2583
 QY 412 -----GluGlnIleValSerLeu-----417
 DB 2584 CAGGCTGAGGCCCAACATGATCGAGAGCTTCGAGAGCTGCAACAAAGGCTCTCCCTCGG 2643
 QY 418 -----GluLysAspThrLeuGlnLeuAspLysGluMetGlu 429
 DB 2644 AGAGCACTCTTAGAACAGAGATTGAAGAAGAGATGTTGGCTTTCAGAAATGAACGACA 2703
 QY 430 GluLysAlaArgSerLeuSerLeuMetGluAlaAlaLeuAlaLysArgMetTyrAsn 449
 DB 2704 GAACGATACGT-----AGCTGCTCGAGCGCCAGCCAGAGAAATTCAGAGCTTTTAC 2757
 QY 450 GlnGlnIleLysArgProArgLeuSerProMetGluMetProValThrSerSerSer 469
 DB 2758 TCTGAAAGCATGAGATTAGGTTTTAGTAACTAGTCTCTT-----TCTAAT 2802
 QY 470 TyrSerProIleTyrArgAspArgSerPheProSerGlnArgAspAspGlnAspGlu 489
 DB 2803 CTCTCCCTCAGGATTCAGCCACAGCTACCCAGGA-----2838
 QY 490 IleSerAlaLeuValSerSerTyrLeuGlyProSerThrSerPheProHisArgSerArg 509
 DB 2839 -----GCTTCTAGCTGGTCTCACAAT-----2859
 QY 510 ArgSerProGluTyrMetValProLeuProHisGlyLeuGlyArgSerValTyrAla 529
 DB 2860 -----CCTACTGGGGGTTTCAGGA-----2877
 QY 530 TyrGluHisLeuAlaProAsnSerTyrSerProGlyHisGlyHisArgLeuHisArgGln 549

Db 2878 -----CCTCACTGGGGTCAATCCCATG----- 2898
Qy 550 TyrSerProSerLeuValHisGlyGlnArgHisProLeuGlnTyrSerPro----- 566
Db 2899 -----GCTGGCACACCAAGCTTGGGTCAATCCGATGCAAGCGGACCCCAACCATGG 2952
Qy 567 -----ProHisGlyGlnGlnLeuProTyrGlyGlnArgValTyrArgHis 584
Db 2953 GGTCACTCCCTCAGGCGCAATCAAGGGGTACCTCGAGTAGCATATAGAGTCCGCAAT 3012
Qy 585 SerProSerGluGluArgTyrLeu-----GlyLeuSerArgGlnA-gSerProArg 601
Db 3013 AGCCCCAGGCTCTGAGCGGACAGCTTCTGGGGACGGACGGAACAGGGGCATGAGCAGA 3072
Qy 602 SerAsnSer 604
Db 3073 AGCAGCAGT 3081

RESULT 2
US-09-723-458-1
Sequence 1, Application US/09723458
Patent No. 6586242
GENERAL INFORMATION:
APPLICANT: Cobb, Melanie
Hutchinson, Michele
Berman, Kevin
Chen, Zhu
TITLE OF INVENTION: TAO PROTEIN KINASES AND METHODS OF USE
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
City: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/723,458
FILING DATE: 27-No. 6586242-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/060,410
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 860098,421
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3312 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 121..3123
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-723-458-1

Alignment Scores:
Pred. No.: 7,198-05 Length: 3312
Score: 151.00 Matches: 150
Percent Similarity: 34.75% Conservative: 95
Best Local Similarity: 21.28% Mismatches: 237

Query Match: 4.86% Indels: 223
DB: 32
US-09-890-475-1 (1-609) x US-09-723-458-1 (1-3312)
Qy 3 AsnTyr-----ProProThrValAlaAlaGlnProThrThrAlaAsn 17
Db 1333 AATTACCAAGAAAGAGATCTTAGAACAAGAGCATCAGCTCCACAGCTTCCA----- 1386
Qy 18 ProLeuLeuGlnArgHisGlnSerGlnArgArgGluLeuProLysIleValGlu 37
Db 1387 CCTCAAGTCTCGTCACAATATCATTATCGTAATAGAGACACTTTGCAACTATACGA 1446
Qy 38 ThrGluSerThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLys 57
Db 1447 ACAGCATCACTG-----GTTACAAGACAGATGCAAGAACATGAG 1485
Qy 58 SerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAsp 77
Db 1486 CAGGACTCTGAATCTAGAGACAGATGCTGTGTTATAAGCGGATGAGCGACAGCAT--- 1542
Qy 78 AspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsn 97
Db 1543 -----CAGACGAGCTGATGACTCTGGAATAATAACTGAAGCGAGAATGACGACAT 1596
Qy 98 GlyVal-----ValLeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeu 113
Db 1597 CGGCTCAGATTAGACAAAGATCTTGAACCTCAGCGCAACAATTCGCTGCAGAAATG--- 1653
Qy 114 SerProArgAsnAsnValSerValGluThrThrValThrValSerGlnProSerGln 133
Db 1654 -----GAGAACTTATTAAAGAAACACCAAGCTTCTATG 1686
Qy 134 GluIleValProGluThrSerAsnLysProGluGlyGlyArgMetCysGluLeuMetCys 153
Db 1687 GAAAGAGAGCTAAAGCTGATGCCAACGAGAG----- 1719
Qy 154 SerLysGlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGlu 173
Db 1720 -----AAAAAATTCACACACACAT-----CAGGCTCAACAGAGAA 1758
Qy 174 GluIleProSerAlaLeuLys-LeuAlaLysGluProAlaLysPheValLeuAspCys11 193
Db 1759 GAACGAAATAGCTTTTGGAGCTCTCAAAAAGAGATATAAACTTCGAAAAGAGCAGCTT 1818
Qy 193 eGlyLysPheTyrLeuGlnGlyArgAlaPheThrLysGluSer-----ProMetSe 211
Db 1819 AAGGAGGAGCTGAATGAAACACAGAGCACACCTAAAGAAAGAAAGCAGGAATGGCTTCA 1878
Qy 211 rSerAlaArgGlnValSerLeuLeuIleLeuGluSerPheLeuLeuMetProAspArg1 231
Db 1879 AAGCAGAGGAGAAATATTCAACATTTTCAGGCAAGAGAGAGAGCTTAATCTTCTGACCT 1938
Qy 231 YLysGlyLysVal-----LysIleGluSerTrpIleLysAspGluAlaGluTh 247
Db 1939 CAAAGGCAGTATCTAGAGCTAGATGCTCGCTTCA-----AAAGAAGATGTTACTT 1992
Qy 247 rAlaAlaValAlaTTPArgLysArgLeuMetThrGlu-GlyGlyLeuAlaAlaGluL 267
Db 1993 GGTCCGCAATAACTTGAACAGGACCTTG-----TCAGGGAGGAGT-----AAACAA 2040
Qy 267 yMetAspAlaArgGlyLeuLeuLeu-----LeuValAlaLac 279
Db 2041 AGCAGACTCAGAGGACCTTAGACATGCAATGTTTACTCGCAGCAGCATCAATCCATCAA 2100
Qy 279 ysPheGlyVal-----ProSerAsnPheArgSerThrAspLeuLeu-AspLeuIleArgMet 297
Db 2101 GAACCTGGAGTTTCGCCACCTCAACACTATTTCAGAGATCGCTGTGAGTGTGATCAGACTG 2160
Qy 298 SerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetVal 317
Db 2161 CAACATCAAACTGAGCTTACTTAACACGCTGGATACATAACAATAGAGAGAGG----- 2208

```

Qy 318 SerGlyIleValGluSerSerIleYsArgGlyMetHisIleGluAlaLeuGluMetVal 337
Db 2209 -----GAACGGAACTAAGACGG-----AAACATGTC----- 2235
Qy 338 TyrThrPheGlyMetGluAspIlePheSerAlaLeuValLeuThrSerPheLeuLys 357
Db 2236 -----ATGGAAGTTCGACAGCAGCCTAAGAGTTTGAGTCTAAAGAACTCCAA 2283
Qy 358 MetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeuAla----- 375
Db 2284 AFAAAAAGCAGTTTCAGGATACCTGCAAAATTCAAACACAGACAGTACAAAGCATTAAGG 2343
Qy 376 -----PheLys 377
Db 2344 AATCACTACTGGAGACTACCAAGAGTGGACAAAGCTGTTCTGAAAGACTCAAG 2403
Qy 378 GluAlaAlaThrLysGlnLeuAlaValLeuSer-----SerValMetGln 392
Db 2404 GAGGAACAGACTCGGAAGTATGCCATCTTGGCTGAGCAGTATGATCATAGCATTAATGAA 2463
Qy 393 CysMetGluThrHis-----LysLeuAspProAlaLysGlu----- 404
Db 2464 ATGCTCTCCACAAAGCTCTGGTGTGGATGAAGCACAGGAGCAGAGATGCCAGGTTTGTG 2523
Qy 405 -----LeuProGlyTyrGlnIleLys----- 411
Db 2524 AAGATCCAGCTACAGCAGGAAGTGGAGCTGTTGATGATGATATCATGACGAAATCAAGATG 2583
Qy 412 -----GluGlnIleValSerLeu----- 417
Db 2584 CAGGCTGAGGCCCAACATGATCGAGAGCTTCGAGAGCTGGAAACAAAGGCTCCCTTCGG 2643
Qy 418 -----GluLysAspThrLeuGlnLeuAspLysGluMetGlu 429
Db 2644 AGAGCACCTTAGACAGAGATTGAAGAGAGATGTTGGCTTTCAGAGTGAAGCAGCA 2703
Qy 430 GluLysAlaArgSerLeuSerLeuMetGluGluAlaAlaLeuAlaLysArgMetTyrAsn 449
Db 2704 GAACGAATACGT-----AGCCTGCTCGAGCCGACGAGAGAAATTCAGAGCTTTTAC 2757
Qy 450 GlnGlnIleLysArgProArgLeuSerProMetGluMetProProValThrSerSerSer 469
Db 2758 TCTGAAGCATGAGATTAGTTTGTATGATGATGATGATGATGATGATGATGATGATGAT 2802
Qy 470 TyrSerProIleTyrArgAspArgSerPheProSerGlnArgAspAspGlnAspGlu 489
Db 2803 CTCTCCCTGAGCATTCAGCCACAGCTACCCAGGA----- 2838
Qy 490 IleSerAlaLeuValSerSerTyrLeuGlyProSerThrSerPheProHisArgSerArg 509
Db 2839 -----GCTTCAGTGTGCTTCACAT----- 2859
Qy 510 ArgSerProGluTyrMetValProLeuProHisGlyGlyLeuGlyArgSerValTyrAla 529
Db 2860 -----CCTACTGGGGGTTTCAGGA----- 2877
Qy 530 TyrGluHisLeuAlaProAsnSerTyrSerProGlyHisGlyHisArgLeuHisArgGln 549
Db 2878 -----CCTACTGGGGGTTTCACAT----- 2898
Qy 550 TyrSerProSerLeuValHisGlyGlnArgHisProLeuGlnTyrSerPro----- 566
Db 2899 -----GGTGGCACACCAAGCTTGGGTTCATCCGATCAAGCGGACCCCAACCATCG 2952
Qy 567 -----ProIleHisGlyGlnGlnLeuProTyrGlyIleGlnArgValTyrArgHis 584
Db 2953 GGTACCCCTTCAGGGCCCAATGCAAGGGTACCTCGAGGTAGGATGATAGGATTCGCAAT 3012
Qy 585 SerProSerGluGluArgTyrLeu-----GlyLeuSerAsnGlnArgSerProArg 601
Db 3013 AGCCCCAGGCTCTGAGCGGACAGCTTCTGGGGGACGCGAGCAGGACAGGCGCATGACGAGA 3072
Qy 602 SerAsnSer 604

```

```

Db 3073 AGCAGGAGT 3081
RESULT 3
US-09-688-188B-20
; Sequence 20, Application US/09688188B
; Patent No. 656715
; GENERAL INFORMATION:
; APPLICANT: FLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0328
; CURRENT APPLICATION NUMBER: US/09/688,188B
; CURRENT FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 09/291,417
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 3824
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-688-188B-20
Alignment Scores:
Pred No.: 0.000164 Length: 3824
Score: 148.50 Matches: 150
Percent Similarity: 34.88% Conservative: 98
Best Local Similarity: 21.10% Mismatches: 227
Query Match: 4.78% Indels: 237
DB: 4 Gaps: 35
US-09-890-475-1 (1-609) x US-09-688-188B-20 (1-3824)
Qy 3 AsnTyr-----ProThrValAlaAlaGlnProThrThrAlaAsn 17
Db 569 AATTACAGAGAAGAGGAGATCTCTAGCAAGAGCATCATCCAACTCCA----- 622
Qy 18 ProLeuLeuGlnArgHisGlnSerGluGlnArgArgGluLeuProLysIleValGlu 37
Db 623 CCCCAAGTATCTCGTCACAAATCACACTATCGTAATCGAGAACACTTGTCTACTATACGG 682
Qy 38 ThrGluSerThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLys 57
Db 683 ACAGCATCACTG-----GTTACGAGGCAATGCAAGAACATGAG 721
Qy 58 SerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAsp 77
Db 722 CAGGACTCTGAGCTTAGAGAACAAATCTCTGGCTATAGCGAATGAGGCGACACAT--- 778
Qy 78 AspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsn 97
Db 779 -----CAAAAGCAACTGATGATGACTCTGGAATAACAGCTAAAGCTGAGTGAACAT 832
Qy 98 GlyValValLeuAlaAlaArgAsnAsnPheHisGlnProMetLeuSerProProArg 117
Db 833 CGCTCTCAGATTA-----GACAAAGATCTTGAACCTCAGCGT 868
Qy 118 AsnAsnValSerValGluThrThrValThrValSerGlnProSerGlnGluIleValPro 137
Db 869 AACAAATTTGTCGAGAAATGAGAAACTTATCAAGAAACACAGGCTGCGATGGAGAAA 928
Qy 138 GluThr-----SerAsnLysProGluGlyGlyArgMetCysGluLeuMetCysSer 154
Db 929 GAGGCTAAGTGTCTCAATGAA-----GAG 955
Qy 155 LysGlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGlu 174
Db 956 AAAAAATTCAGCAACATATT-----CAGGCCCAACAGAGAAAGAA 997

```

```
QY 175 IleProSerAlaLeu-----lysLeuAlaLysGluProAlaLys 167
   |||
   |||
   |||
Db 998 CTGAATAGTTTCTCGAGTCCCGAAGAGAGAGTATAAATTCGAAAGAGACAGCTTTAAA 1057
   |||
   |||
   |||
QY 188 PheValLeuAspCysIleGlyPheThrLeuGlnGlyArg-----AlaPhe 204
   |||
   |||
   |||
Db 1058 GAGGAGCTAAAT--GAAACCAGAG-TACCCCAAAAGAAAGAAACAGAGGAGTGCCTTC 1113
   |||
   |||
   |||
QY 205 ThrLysGluSer-----ProMetSerSer 212
   |||
   |||
   |||
Db 1114 AAAGCAGAAGGAGAGATATACAGCATTTTCCAAACGAGAGAGAGTAACTTCTTCGAGC 1173
   |||
   |||
   |||
QY 213 AlaArgGlnValSerLeuLeuIleLeuGluSerPheLeuMetProAspArgGlyLys 232
   |||
   |||
   |||
Db 1174 TCAGACAAAT-----CCTAGAGCTGGAATG 1200
   |||
   |||
   |||
QY 233 GlyLysValLysIleGluSerTrpIleLysAspGluAlaGluThrAlaAlaValAlaTrp 252
   |||
   |||
   |||
Db 1201 CCGTCGCTTCAAGAGAGAAAT-----GTTACTTGG 1230
   |||
   |||
   |||
QY 253 ArgLysArgLeuMetThr-----GluGlyGlyLeuAlaAlaAlaGluLysMetAsp 269
   |||
   |||
   |||
Db 1231 CGCTCATAACTTAGAGCAGGACCTTGTCAGGAGAGAGTT-----AAACAAAGACAGAC 1284
   |||
   |||
   |||
QY 270 AlaArgGlyLeuLeuLeuValAlaCysPheGlyValProSer----- 284
   |||
   |||
   |||
Db 1285 TCAGAAGGACTTAGA-----GCATGCCATGCTACTCCGACAGCATGAATCTATGCA 1335
   |||
   |||
   |||
QY 285 -----AsnPheArgSerThrAspLeuLeu-AspLeuIleArgMe 297
   |||
   |||
   |||
Db 1336 AGAACTGAGTTCCGCCACCTCAACAAATTCAGAGATGCGCTGTGAGTTGATCAGATT 1395
   |||
   |||
   |||
QY 297 tSerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetVa 317
   |||
   |||
   |||
Db 1396 ACAGCATCAACTGAGCTCACTACCAGCTGGATATATTAAGCGAAGA----- 1444
   |||
   |||
   |||
QY 317 lSerGlyIleValGluSerSerIleLysArgGlyMethIleGluAlaLeuGluMetVa 337
   |||
   |||
   |||
Db 1445 -----GAAAGAGAACTTAAGACGA---AAGCATGTC----- 1471
   |||
   |||
   |||
QY 337 lTyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLys 357
   |||
   |||
   |||
Db 1472 -----ATGGAAGTTCCGACACACCTTAAGAGTTTGAAGTCTAAAGAACTCCA 1518
   |||
   |||
   |||
QY 357 sMetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeuAla----- 375
   |||
   |||
   |||
Db 1519 AATRAAAAGCAGTTTCAGGATACCTGCAAAATCCAAACACAGACAGTCAAAAGCATTAAAG 1578
   |||
   |||
   |||
QY 376 -----PheLys 377
   |||
   |||
   |||
Db 1579 AATCACCCTGCTGGAGACTACACCAAGAGAGTGAGCACAAAGCTGTTCTGAAACGGCTCAA 1638
   |||
   |||
   |||
QY 377 sGluAlaAlaThrLysGlnLeuAlaValLeuSer-----SerValMetG1 392
   |||
   |||
   |||
Db 1639 GGAGGAACAGACCCCGGAATTAGCTATCTTGCTGAGCAGATGATGATCACACCATTAATGA 1698
   |||
   |||
   |||
QY 392 nCysMetGluThrHis-----LysLeuAspProAlaLysGluLeuProGlyTrpGlnI1 410
   |||
   |||
   |||
Db 1699 AATGCTCTCCACACAAAGCCCTCGGTTTGGATTGAAGACAGCAAGGAGAGTCCCGAGTTT 1758
   |||
   |||
   |||
QY 410 eLysGluGlnIle-----ValSerLe 417
   |||
   |||
   |||
Db 1759 GAAGATGAGCTGTCAGCAGGAACTGGAGCTGTTGATGCGTATCAGAGCAAAATCAAGAT 1818
   |||
   |||
   |||
QY 417 uGluLysAspThrLeuGlnLeuAspLysGluMetGluGlyLysAlaArgSerLeuSerLe 437
   |||
   |||
   |||
Db 1819 GCAAGCTGAGGCA---CAACATGATCGAGAGCTTCGCGAGCTTGAACAGAGGAGTCCCT 1875
   |||
   |||
   |||
QY 437 u-----MetGluGluAlaAlaLeuAlaLysArgMetTyAs 449
   |||
   |||
   |||
Db 1876 CCGGAGGGCAGCTCTTAGAACAAAGAGATTGAAGAGAGATGTGGCT-----TTGCGAAG 1929
   |||
   |||
   |||
QY 449 nGlnGlnIleLysArgProArgLeuSerProMetGluMetProProValThrSerSerSe 469
```

```
Db 1930 TGAGCCACAGAACGCAATACGA----- 1951
   |||
   |||
   |||
QY 469 rTySerProIleTyArgAspArgSerPheProSerGlnArgAspAspGlnAspG1 489
   |||
   |||
   |||
Db 1952 -----AGCCTGTTGGAACGTCACAGCCAGAGA 1977
   |||
   |||
   |||
QY 489 uIleSerAlaLeuValSerSerTyLeuGlyProSerPheProHisArgSerAr 509
   |||
   |||
   |||
Db 1978 GATTGAAGCTTTTGACTCTGAAAGCATGAGACTAGGTTTTAGATATATGCTCTTCTTAA 2037
   |||
   |||
   |||
QY 509 gArgSerProGluTyMetValProLeuProHisGlyGlyLeuGlyArgSerValTyAl 529
   |||
   |||
   |||
Db 2038 TCTCTCCCTGAG-----GC 2052
   |||
   |||
   |||
QY 529 aTyGluHisLeuAlaProAsn-----SerTySerProGlyHisGly----- 543
   |||
   |||
   |||
Db 2053 ATTACGCCACAGTACCCGGGAGCTTCTGTTGGTCACACAACCTACTGGGGGTCACAG 2112
   |||
   |||
   |||
QY 544 ---HisArgLeuHisArgGlnTySerProSerLeuValHisGlyGlnArgHisProLe 562
   |||
   |||
   |||
Db 2113 ACCTCACTGGGTCTCCATGGTGGCCCAACACCAAGCTTGGGGC-----CATCCAAT 2166
   |||
   |||
   |||
QY 562 uGlnTySerPro-----ProIleHisGlyGlnGlnLeuProTyG1 577
   |||
   |||
   |||
Db 2167 GCAAGGTGAGCCACAGCATGGGTGCGCCCAATCAAGGGGTACTCTCGAG 2226
   |||
   |||
   |||
QY 577 yIleGlnArgValTyArgHisSerProSerGluGluArgTyLeu-----GlyLe 594
   |||
   |||
   |||
Db 2227 TAGCATATGGAGTCCGCAATAGACCCCAAGCTCTGAGCGGACAGCTTCTGGGGGACG 2286
   |||
   |||
   |||
QY 594 uSerAsnGlnArgSerProArgSerAsnSer 604
   |||
   |||
   |||
Db 2287 GAACGAGCGGCGCATGAGCAGACGAGT 2317
   |||
   |||
   |||
RESULT 4
US-09-291-417D-20
; Sequence 20, Application US/09291417D
; Patent No. 6680170
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHITE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0329
; CURRENT APPLICATION NUMBER: US/09/291,417D
; CURRENT FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 20
; LENGTH: 3824
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-291-417D-20
Alignment Scores:
Pred. No.: 0.000164 Length: 3824
Score: 148.50 Matches: 150
Percent Similarity: 34.88% Conservative: 98
Best Local Similarity: 21.10% Mismatches: 227
Query Match: 4.78% Indels: 237
DB: 4 Gaps: 35
US-09-890-475-1 (1-609) x US-09-291-417D-20 (1-3824)
QY 3 AsnTyT-----ProProThrValAlaAlaGlnProThrThrAlaAsn 17
   |||
   |||
   |||
Db 569 AATTACAGAAAGAGGAGATCTTAGAACAGAGATCAGATCCCAATCTCCA----- 622
   |||
   |||
   |||
QY 18 ProLeuLeuGlnArgHisGlnSerGluGlnArgArgGluLeuProLysIleValGlu 37
   |||
   |||
   |||
```

623 CCCAAGTATCTCGTCACAAATCACTATCGTAAATCGAGAACACTTTGCTACTATACGG 682
Db
38 ThrGluSerThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLys 57
Qy
683 ACAGCATCACTG-----GTTACGAGGCAAAATGCAAGACATGAG 721
Db
58 SerIleAspGluLeuAlaPheSerValAlaValGluThrPheLysArgGlnPheAsp 77
Qy
722 CAGGACTCTGAGCTTAGAGAACAAATGCTCGCTATTAAGCGAATGAGCGGACACAT--- 778
Db
78 AspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsn 97
Qy
779 -----CAAAAGCACTGATGCTCTGAAACCAAGCTAAAGCTGAGATGATGAACAT 832
Db
98 GlyValValLeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerProArg 117
Qy
833 CGCCTCAGATTA-----GACAAAGATCTTGAACATCAGCGT 868
Db
118 AsnAsnValSerValGluThrThrValThrValSerGlnProSerGlnGluIleValPro 137
Qy
869 AACAAATTTGCTGCAAAATGAGAACTTATCAAGAAACACAGGCTGCCATGAGAAA 928
Db
138 GluThr-----SerAsnLysProGluGlyArgMetCysGluLeuMetCysSer 154
Qy
929 GAGGCTAAAGTATGATCCCAATGAA-----GAG 955
Db
155 LysGlyLeuArgLysIleThrAlaAsnIleSerAspGlnAlaLysLeuMetGluGlu 174
Qy
956 AAAAAATTCACAAATATT-----CAGGCCACACAGAAAGAA 997
Db
175 IleProSerAlaLeu-----LysLeuAlaLysGluProAlaLys 187
Qy
998 CTGAATAGTTTCTCGAGTCCAGAAAAGAGAGATATAAACTTCGAAAAGACGACTTAA 1057
Db
188 PheValLeuAspCysIleGlyLysPheThrLeuGlnGlyArgArg-----AlaPhe 204
Qy
1058 GAGGAGCTTAAAT---GAAAAACAGAG-TACCCCCAAAAGAAAAGAGAGGTGCTTTC 1113
Db
205 ThrLysGluSer-----ProMetSerSer 212
Qy
1114 AAGCAGAGGAGGAGATATACAGCATTTCCACGACAGAGAGAGTAACTTCTTCGACG 1173
Db
213 AlaArgGlnValSerLeuLeuIleLeuGluSerPheLeuLeuMetProAspArgGlyLys 232
Qy
1174 TCAAAAGACAATA-----CCTAGAGCTGGAATG 1200
Db
233 GlyLysValLysIleGluSerTrpIleLysAspGluAlaGluThrAlaAlaValAlaTrp 252
Qy
1201 CCGTCGCTTCAAGAGAGAT-----GTTACTTGG 1230
Db
253 ArgLysArgLeuMetThr-----GluGlyGlyLeuAlaAlaGluLysMetAsp 269
Qy
1231 CGGTATTAATCTAGAGCAGGACCTTGTCTAGGAGGAGTT-----AAACAAAAGACAG 1284
Db
270 AlaArgGlyLeuLeuLeuValAlaCysPheGlyValProSer----- 284
Qy
1285 TCAGAAGGACTTAGA-----GCATGCCATGTCTCTCCAGCAGCATGATCTATGCA 1335
Db
285 -----AsnPheArgSerThrAspLeuLeu-AspLeuIleArgMe 297
Qy
1336 AGAATCGGAGTTCCGCCACCTCAACACAAATTCAGAAGATGCGCTGTGATGATCAGATT 1395
Db
297 tSerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetVa 317
Qy
1396 ACAGCATCAAACTGACTCACTAACCGCTGGAATATATAAGCGAGA----- 1444
Db
317 lSerGlyIleValGluSerSerIleLysArgGlyMethHisIleGluAlaLeuGluMetVa 337
Qy
1445 -----GAACGAGAACTAAGACGA--AAGCATGTG- 1471
Db
337 lTyThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLys 357
Qy
1472 -----ATGGAAGTTCGACAAACAGCGCTAAGATTTGAAGTCTTAAAGAACTCCA 1518
Db

357 sMetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeuAla----- 375
Qy
1519 AATAAAAAAGCAGTTTTCAGGATACCTGCAAAATCCAAACCCAGACAGTAGTCAAAAGCATTAG 1578
Db
376 -----PheLys 377
Qy
1579 AAATCACCTGCTGGAGACTACACCAAGAGTGAGCACAAAGCTGTTCTGAAACGGCTCAA 1638
Db
377 sGluAlaAlaThrLysGlnLeuAlaValLeuSer-----SerValMetG1 392
Qy
1639 GGAGGAACAGACCCCGAAATTAGCTATCTTGCTGAGCAGTATGATCAACAGCATTAATGA 1698
Db
392 nCysMetGluThrHis-----LysLeuAspProAlaLysGluLeuProGlyTrpGlnI1 410
Qy
1699 AATGCTCTCCACACAGCCCTGCGTTTGATGAAGACACAGGAGCAGAGTGCAGGTTT 1758
Db
410 eLysGluGlnIle-----ValSerLe 417
Qy
1759 GAAGATGACGCTGCAGCAGGAACCTGGAGCTGTTGAATGCGTATCAGAGCAAAATCAAGAT 1818
Db
417 uGluLysAspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLe 437
Qy
1819 GCAAGCTGAGGCA---CAACATGATCGAGCTTTCGCGAGCTTGAACAGAGGTTCTCCT 1875
Db
437 u-----MetGluGluAlaAlaLeuAlaLysArgMetTyrAs 449
Qy
1876 CCGGAGGCGACTCTTAGAAACAAAGATTGAAGAGAGATGTTGGCT-----TTGCAGAA 1929
Db
449 nGlnGlnIleLysArgProArgLeuSerProMetGluMetProProValThrSerSerSe 469
Qy
1930 TGAGCCCGCAGAACGAATACGA----- 1951
Db
469 rTyrSerProIleTyrArgAspArgSerPheProSerGlnArgAspAspGlnAspG1 489
Qy
1952 -----AGCCTGTTTGAACGCTCAAGCCAGAGA 1977
Db
489 uIleSerAlaLeuValSerSerTyrLeuGlyProSerThrSerPheProHisArgSerAr 509
Qy
1978 GATTGAAGCTTTGACTCTGAAAGCATGAGACTAGAGTATTAATATATGCTCTTCTTAA 2037
Db
509 gArgSerProGluTyrMetValProLeuProHisGlyGlyLeuGlyArgSerValTyrAl 529
Qy
2038 TCTCTCCCTCGAG-----GC 2052
Db
529 aTyrGluHisLeuAlaProAsn-----SerTyrSerProGlyHisGly----- 543
Qy
2053 ATTCAAGCCACAGCTACCCGAGGAGCTTCTGTTGGTCAACAACCTTCTGGGGGTCCAGG 2112
Db
544 ---HisArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArgHisProLe 562
Qy
2113 ACCTCATGGGGTCATCCCATGGGTGGGCCACCAACAGCTTGGGCG-----CATCCAT 2166
Db
562 uGlnTyrSerPro-----ProIleHisGlyGlnGlnGlnLeuProTyrG1 577
Qy
2167 GCAAGGTGGACCCCGAGCCATGGGGTCACTTCCAGGGCCAATGCAAGGGGTACCTCGAG 2226
Db
577 yIleGlnArgValTyrArgHisSerProSerGluGluArgTyrLeu-----GlyLe 594
Qy
2227 TAGCAGTATGGAGTCCGCAATAGCCCCCAGCGCTCTGAGCGGACAGCTTCTGGGGGACG 2286
Db
594 uSerAsnGlnArgSerProArgSerAsnSer 604
Qy
2287 GACGAGCAGGGCATGAGCAGAGACGAGT 2317
Db

RESULT 5

US-09-491-356C-7

; Sequence 7, Application US/09491356C

; Patent No. 6566061

; GENERAL INFORMATION:

; APPLICANT: Philibert, Robert A.

; APPLICANT: Gibbs, Edward I.

; APPLICANT: Delisi, Lynn

; TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCTG4 REGION OF XOL3

; FILE REFERENCE: 9465.6US11
; CURRENT APPLICATION NUMBER: US/09/491.356C
; PCT FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: PCT/US99/09365
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 60/083,465
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 6558
; TYPE: DNA
; ORGANISM: Mus musculus
; US-09-491-356C-7

Alignment Scores:

Pred. No.:	0.00535	Length:	6558
Score:	137.50	Matches:	137
Percent Similarity:	33.46%	Conservative:	118
Best Local Similarity:	17.98%	Mismatches:	226
Query Match:	4.43%	Indels:	281
DB:	4	Gaps:	36

US-09-890-475-1 (1-609) x US-09-491-356C-7 (1-6558)

Qy	17	AsnProLeuLeuGlnArgHisGlnSerGluGlnArgargGluLeuProLysIleVal	36
Db	3928	GACCAGTGTTCAGTAGTGCACGAGCCGAGCCG	3975
Qy	37	GluThrGluSerThrSerMetAspIleThrIleGlnSerLysGlnProGln	54
Db	3976	TGCTACCCACATCGACTGCTGACATCAGGATGAGAAACCCGAGCGCGCAT	4035
Qy	55	-----PheLeuLysSerIleAspGluLeuAlaPheSerValAlaValGluThrPhe	72
Db	4036	AAACGTATTTCTCAAGATTTTAGCAATGG	4071
Qy	73	LysArgGlnPheAspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSer	92
Db	4072	CGCAGTCTCTTTGGAGTACAGTGTGATGATCAGCAGACCCCAACACT	4122
Qy	93	LysLeuGluSerAsnGlyValValLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAla	112
Db	4123	-----GAGATGAATCTCTCTTG	4140
Qy	113	LeuSerProArgAsnAsnValSerValGluThrThrValThrValSerGlnProSer	132
Db	4141	-----GAGAACATTTGCC	4182
Qy	133	GlnGluIleValProGluThr	148
Db	4183	CGACGACGGGTCTCTCTGGAAGTACAGCAACATGCGCCAGCAGCAGCAGC	4242
Qy	149	CysGluLeuMetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGln	168
Db	4243	AAACCTGTGCTCAGCTCT	4287
Qy	169	AlaLysLeuMetGluGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLysPhe	188
Db	4288	GCTCTCTCATGTGCACAACTGCCACITTCAGTCCAG	4329
Qy	189	ValLeuAspCysIleGlyLysPheTyrLeuGlnGlyArg	205
Db	4330	GTATTAAGTGTCTGGGAGAACTGGAAAGGCTCAGCACCTGGGCTCTCTTCCCGT	4389
Qy	206	LysGluSerProMetSerSerAlaArgGlnValSerLeuLeuIleLeuGluSerPheLeu	225
Db	4390	AAAGAACGAGATCGACAAACAGACAGATGTCCTGTTTGACCCACAGCCTTTCTTA	4449
Qy	226	-----LeuMetProAspArgLysGlyLysValLysIleGluSerTrp	240
Db	4450	TCACTGCTGCTAAACATGCTGAAAGGACAGGATGAGCAGCGCGGAGCTCTCTGCTCC	4509

Qy	241	IleLysAspGluAlaGluThrAlaAlaValAlaTrpArg	253
Db	4510	CTCCACAGCCAGGTGCACGAGATTGTGATTAAATTGGCGAGAAACACGACTTAGATGAT	4569
Qy	254	-----LysArgLeuMetThrGlu	261
Db	4570	TGCAAAACCAAGCAGCTTAATCATGAGGACCTCAAACTCGGCTCAACTGGTGGGGGC	4629
Qy	262	Leu	275
Db	4630	ATGTTTGACACTGTGCAGCGTAGTACCAGCAGACTACAGAGTGGCGCCAGCTTCTCCT	4689
Qy	276	LeuValAlaCysPheGlyVal	291
Db	4690	GAGATCATCATCAGCGGCACTGTGACATGCGAGTCTAAATGAGCTCTTACTACTGTG	4749
Qy	292	LeuAspLeuIleArgMetSerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGln	311
Db	4750	TTGGACATGCTGAGCGTCTTATCAACGGAACGTTGGCTGCAGACATG	4797
Qy	312	PheLeuValProMetValSerGlyIleValGluSerSerIleLysArgGlyMethIle	331
Db	4798	-----TCCAGTATCTCGCAAGGC	4815
Qy	332	GluAlaLeuGluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuVal	351
Db	4816	-----AGCATGGAGGAAACCAACGCTGCATATATGAAC	4848
Qy	352	LeuThrSerPheLeu	366
Db	4849	CTGGTGAAGAGCTTCAGAAAGACTTGGGGGAGCGCAATCAGACAGTCTGAGAAAGTT	4908
Qy	367	LysArgLysAlaGlnSerProLeuAlaPheLysGluAlaAlaThr	381
Db	4909	CACCAACTCTGCCACTCCCAAGCAGAACCCAGATGTCAACTGTGAGCCACAGGCG	4968
Qy	382	-----LysGlnLeuAlaValLeuSerSerVal	390
Db	4969	TCCTTTATGACACCAAGGGCAACAAGATTGTGGCTTCGATTCCATCTTCAAGAAGGAG	5028
Qy	391	-----MetGlnCysMetGluThrHisLysLeuAspProAlaLysGluLeuProLysGln	409
Db	5029	GGTCTACAGTTTCTACCAACAAAGATCTCTCCC	5070
Qy	410	IleLysGluGlnIleValSerLeuGluLys	422
Db	5071	CTTTTGGGGCTGAGCCATCAACAGCACCCTGTCATGGCGCTGGTTTGGCAGCTC	5130
Qy	423	GlnLeuAspLysGluMet	439
Db	5131	CGAGTGCACGAGAGTGGCAGGGGAGGAGCAGCAGCGCTGTG	5178
Qy	440	GluAlaAlaLeuAlaLysArgMetTyrAsnGlnIleLys	456
Db	5179	-----CTCTATCATACCCCTGAGGCTTCGACCCAGACCTAT	5217
Qy	457	-----LeuSerProMetGluMetProProValThrSerSerSerTyrSerProLys	474
Db	5218	TACCTGGAACCACTACCTCTGCCCCCAGAAAGATGAGGACCCACAGCCCTCCCTACTA	5277
Qy	474		474
Db	5278	GAGCTGAGAAAAAGGCTCTCTGAGCCCCCAAGACTGACAAACAGGGGCTCTCTCTCG	5337
Qy	475	-----ArgAspArgSerPheProSerGln	482
Db	5338	AGCACTGAGCGCGCAAAAAGAGTCTACCAAGGGCAAAAACAGCAGCCAGCAGCCACC	5397
Qy	483	ArgAspAspGlnAspGluIleSerAlaLeuValSerTyrLeuGlyProSerThr	502
Db	5398	AAGAACGAGGACTATGGC	5430

9991 GCTGTGCTCCGCGCT-----GCAAAATCGGCACCTGCAGTTTTCGCG 10032

131 ProSerGlnGluIleValProGluThrSerAsnLysProGluGlyGlyArgMetCysGlu 150

10033 AAATCGTTAACTTGGCGCCTCGGCCATCCATAAAACAACAAGAACCAACAGCAAGATG 10092

151 -LeuMetCysSerLysGlyLeuArgLysTyrIle-----TyrAlaAsnIleSerAspG1 168

10093 GATCTTCTGTTCGGGAACGCAATCGCCCATGTCCACCGATACCCACGCCCTCGCTAGCG 10152

168 nAlaLysLeuMetGlnGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLys-P 188

10153 CTCGCCAAGCCCGCGCTTCGCGCGCTGCCCTTCGCCTTCGCCAAACGCCACGCGCTGC 10212

188 heValLeuAspCysIleGlyLysPheTyrLeuGlnGlyArgArgAlaPheThrLysGluS 208

10213 TGTGTGCGAGCCCTTCGCGCCAGGTCACGCTCAGGTGCGCGCGGT----- 10259

208 erProMetSerSerAlaArgGlnValSerLeuLeuIleLeuGluSerPheLeuLeuMetP 228

10260 -----GCCAGCTCGCGCGCTGCAGGAG-----G 10284

228 roAspArgGlyLysGlyLysValLysIleGluSerTrpIleLysAspGluLa-----G 246

10285 CCCAGCGCTTCGCGCGCGCTGCTCCGCTGCATCTGCTGAGCGCCGAGGCTTCGAGC 10344

246 LuThrAlaAlaValAlaTrpArgLys-----ArgLeuMetThrGluGlyG 261

10345 AGGAGCTGGCCCTTGGCCTACCAAGCGGCACTCTCCAGGTGCGGAGATGCCGAGGGC- 10403

261 lyLeuAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuLeuValAlaCysPheG 281

10404 -----ATGGGTGCGGAACCTGACCTAGCCAGCGCTGCCGAACTCACTCCC----- 10448

281 lyValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySerA 301

10449 -----GAATCGCGGACCTGCTCGAG----- 10469

301 snGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIleV 321

10470 -----CAGGAAGATGACGCGCGCATCATCGCTGATCAACGCCATCC 10512

321 alGluSerSerIleLysArgGly-----MetHisIleGluAlaLeuGlu-----M 336

10513 TCAGCAGGCGGATCAAGCGCGCGCTCCGACATCCACTCGAAACCTTCAGAAAACGCC 10572

336 etValTyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheL 356

10573 TGTGTGTGCGCTTTCGCTCGACGCG-----ATCCTCGCGGAGTGA 10614

356 euLysMetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeuAlaP 376

10615 TCGAAACGCGCGCGCGAG-----CTGGCG- 10637

376 heLysGluAlaAlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluT 396

10638 -----GCCTGCTGTCTCGCGGTCAAGTTCATGGCG- 10670

396 hrHisLysLeuAspProAla--LysGluLeuPro----- 406

10671 -----CGCTGGACATCGCGGAGAGCGGTACCCAGGAGCGCGCTATTTCTGCTCAAG 10725

407 ----GlyTrpGlnIleLysGluGlnIleValSerLeu-----GluLysA 420

10726 TCGCGCGTTCGAGGTGGATATCCGCTTCACCTCGCTCGGCCAACGCGAGCGGG 10785

420 spThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeu----- 437

10786 TGTGTGTGCGCTG-----CTGCACAAGAGCGCGCGCGCTGTCTGCTCAGCATC 10836

438 -----MetGluGluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgPro- 455

10837 TGGGCATGAGCAGCGC-----GACCGCGCGCTGTCTGACGACAACTCTGGCAGCGCG 10890

RESULT 8

US-09/479-453-29
Sequence 29, Application US/09479453
Patent No. 6313283
GENERAL INFORMATION:
APPLICANT: Gerritse, Gijssbert
APPLICANT: Quax, Wilhelmus J.
TITLE OF INVENTION: EXPRESSION SYSTEM FOR ALTERED
TITLE OF INVENTION: EXPRESSION LEVELS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genencor International
STREET: 925 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1013
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/479,453
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/911,853
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Glaister, Debra J
REGISTRATION NUMBER: 33,888
REFERENCE/DOCKET NUMBER: GC361-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-846-7620
TELEFAX: 650-845-6504
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 17612 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

```

; TOPOLOGY: linear
US-09-479-453-29
Alignment Scores:
Pred. No.: 0.0803 Length: 17612
Score: 133.00 Matches: 131
Percent Similarity: 35.95% Conservative: 98
Best Local Similarity: 20.57% Mismatches: 224
Query Match: 4.28% Indels: 184
DB: 4 Gaps: 31

US-09-890-475-1 (1-609) x US-09-479-453-29 (1-17612)
QY 14 ThrThrAlaAsnProLeuLeuGlnArgHisGlnSerGluGlnArgArgGluLeuPro 33
Db 9754 ACCTCGATCTCGCGCTGACCAATCGTAGCTACCCCAACGACGCGCGGTG 9813
QY 34 LysileValGluThrGluSerThrSerMetAspIleThrileGlyGlnSerLysGlnPro 53
Db 9814 AAGATCGCGAGATGGAGCTGACCTCGCGTTCATCGCC----- 9852
QY 54 GlnPheLeuLysSerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLys 73
Db 9853 -----TCGGCCTTCGGCGCAACACGCGGTGTTCCTCAAG 9885
QY 74 ArgGlnPheAspAspLeuGlnLysHisIleGluSerIleGluAenAlaIleAspSerLys 93
Db 9886 CAC-----CAGCGTCACGAAGACAAATAAGGGTCATCCCTTGCTGAACAG 9930
QY 94 LeuGluSerAsnGlyValValLeuAlaAlaArgAsnAsnAsn-----PheHisGln 110
Db 9931 CCCCGCGCCGTGCGCGGCGCTTTTGTGTCACGCGCTTACGTCATCACTTCTGCGCCAG 9990
QY 111 ProMetLeuSerProProArgAsnAsnValSerValGluThrThrValThrValSerGln 130
Db 9991 GCTGTGTCGCGCT-----GCAAAATCGGCACATGCAGTTTTCGCGC 10032
QY 131 ProSerGlnGluValProGluThrSerAsnLysProGluGlyArgMetCysGlu 150
Db 10033 AAATCCGTAACTTGGCGCCTCGCGCCATGCCATAAAACAAACAAACAGCAGATG 10092
QY 151 LeuMetCysSerLysGlyLeuArgLysTyrlle-----TyrAlaAsnIleSerAspG 168
Db 10093 GATCTTCTGTTCGGGGAACGATCCGCCCATGTCCACGATACCACCGCGCGCTGACGG 10152
QY 168 nAlaLysLeuMetGluGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLys-P 188
Db 10153 CTCGGCAAGCCCGCTTGGCGCGCTGCGCTTGCCTTCGCCATACGCGCGGTGC 10212
QY 188 heValLeuAspCysIleGlyLysPheTyrlleGlnGlyArgArgAlaPheThrLysGluS 208
Db 10213 TGCTGCGGAGCCCTTCGGCGAGTCCAGTGTGAGTGGCGCGGT----- 10259
QY 208 erProMetSerSerAlaArgGlnValSerLeuLeuIleLeuGluSerPheLeuLeuMetP 228
Db 10260 -----CCGAGCTGGCGCGGTGCAGGAG-----G 10284
QY 228 roAspArgGlyLysGlyLysValLysIleGluSerTrpIleLysAspGluAla-----G 246
Db 10285 CCCAGCGCTTCGGCGCGCGGTGCTGCGCTGCATCTGCTGGAGCCGAGGCTTCGAGC 10344
QY 246 luThrAlaAlaValAlaTrpArgLys-----ArgLeuMetThrGluGlyG 261
Db 10345 AGGAGCTGGCGCTGCGCTTACGAGCGGACCTCTCCGAGGTGCGGACAGATGCGCGGCGC- 10403
QY 261 lyLeuAlaAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaLysPheG 281
Db 10404 -----ATGGGTGCCAACTTCACCTAGCAGCCTGGCGCACTCACCTCCC----- 10448
QY 281 lyValProSerAsnPheArgSerThrAspLeuLeuAspLeuLeuLeuLeuSerGlySera 301
Db 10449 -----GAATCCGCGACCTGCTGGAG----- 10469

RESULT 9
US-08-977-171-1
; Sequence 1, Application US/08977171
; Patent No. 6232112
; GENERAL INFORMATION:
```

APPLICANT: CATCHESIDE, DAVID E.
 TITLE OF INVENTION: REAGENTS AND METHODS FOR DIVERSIFICATION
 TITLE OF INVENTION: OF DNA
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
 STREET: 3100 No. 6232112west Center, 90 South 7th Street
 CITY: Minneapolis
 STATE: MN
 COUNTRY: USA
 ZIP: 55402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/977,171
 FILING DATE: 24-NOV-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Skoog, Mark T
 REGISTRATION NUMBER: 40,178
 REFERENCE/DOCKET NUMBER: 10552.13US01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 612-332-5300
 TELEFAX: 612-332-9081
 TELEX:

QY 523 uGlyArgSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSerProGlyHisG1 543
Db 2413 TTCCCGC-----CATGG 2424
QY 543 yHisArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArgHisProLeuG1 563
Db 2425 TCGCGCGCTG-----CNAAGATGTGTTCG 2451
QY 563 nTyrSerProPro-----lleHisGlyGlnGlnLeuPro----- 575
Db 2452 CCTCTCCCGCGCGCGAGAACCATCACTCCCGAGATTGTCCAGCTCGCTCACAAGG 2511
QY 576 -TyrGly---lleGlnArgValTyrArgHisSerProSer----- 587
Db 2512 TTGGGCGCGAGTCCATCGTCTCCCGCGGTGCCAGCGGTAGCTGCCATGCGCTACG 2571
QY 588 ---GluGluArgTyrLeuGlyLeuSerAsnGlnArgSerProArgSerAsnSerLe 606
Db 2572 GCACCGAGAGCATCACCAAGTGCACAAGATTCTCGGCC-CCGTAACCAAGTTCGTCACT 2630
QY 506 uAspProLys 609
Db 2631 GCTGCCAAGA 2640

RESULT 10

US-09-620-312D-372
; Sequence 372, Application US/09620312D
; Patent No. 6569662

GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunwei
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784C1P2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt-fl_genes Version 1.0
; SEQ ID NO 372
; LENGTH: 7393
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (119)..(3012)
US-09-620-312D-372

Alignment Scores:

Pred. No.:	0.0337	Length:	7393
Score:	130.50	Matches:	137
Percent Similarity:	30.80%	Conservative:	86
Best Local Similarity:	18.92%	Mismatches:	274
Query Match:	4.20%	Indels:	227
DB:	4	Gaps:	29

US-09-890-475-1 (1-609) x US-09-620-312D-372 (1-7393)
QY 1 MetSerAsnTyrPro-----ProThrValalaalaGlnPro 12
Db 859 ATGCCCAACAGTCCCAACTCAAGCCCATTCAGCCCAAGCCCACTGTTATGGAGAACCT 918
QY 13 ThrThrThrAlaAsnProLeuLeu-----GlnArgHisGlnSerGluGlnArg 28
Db 919 ---TTCAACAGTCAACCCCTGCCTTGACTCCAGCCAGGCAAGAAAGAAAGAAAGAAAA 975
QY 29 ArgArgGluLeuProLysIleValGluThrGluSerThrSerMetAspIleThrIleGly 48
Db 976 AAGAAGGATCTTCAAGGAACCTTGAAGTCTCTGACCCCTGGGAAGGTGTGTGACGA 1035
QY 49 GlnSerLysGlnProGlnPheLeuLysSerIleAspGluLeuAlaAlaPheSerValala 68
Db 1036 GAGGAAGGCAAAAGGCCATTC----- 1056
QY 69 ValGluThrPheLysArgGlnPheAspLeuGlnLysHisIleGluSerIleGluAsn 88
Db 1057 -----AGGAATCTTCAGGAATGGATGAAATGAGGGGCTCTTAAT 1101
QY 89 AlaIleAspSerLysLeuGluSerAsnGlyValValLeuAlaAlaArgAsnAsnAsnPhe 108
Db 1102 GGCTCATCAGACCCCAAGCCGACTGGCTAGCATCAAGGCTGAAGCCGACAAGATC 1161
QY 109 HisGlnProMetLeuSerProArgAsnAsnValSerValGluThrThrValThrVal 128
Db 1162 TACAGTTTCAGCAATGCCCCAGCCCTTCCATTGGAGGAGTAGCGGCTCTGAAAC 1221
QY 129 SerGlnProSerGlnGluIleValPro-----GluThrSerAsnLysProGluGly 145
Db 1222 ACTACCCCTACTCAGCCCTGACTCCCTTACATGTGTGACCCAGACATGAGGCTGAAGCC 1281
QY 146 GlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsnIle 165
Db 1282 AGC-----TCAGTCAAAACCAACAGCCCTGTCATCTCTGACATC 1320
QY 166 SerAspGlnAlaLys-----LeuMetGluGluIleProSer-----AlaLeu 179
Db 1321 TCTGATGTGGGAGGATGGGAGGGCAAGGTAGACAGTGTCAATCAAGGACGCCGNA 1380
QY 180 LysLeuAlaLysGluProAlaLysPheValLeu-----AspCysIleGly 194
Db 1381 CAGTTGGTTAAAGAGGGGCTAAGAAAACTCTTTTCCCTCCCTCAGCCTCAGACAAAGAC 1440
QY 195 LysPheTyrLeuGlnGlyArgArgAlaPheThrLysGluSer----- 208
Db 1441 TCACCATATTACCAAGGCTTTGAGAGTTTACTATTCTCCAAGTTATGCACAGTCCAGCCCT 1500
QY 209 -----ProMetSerSerAlaArgGlnValSerLeuLeuIleLeuGluSerPhe 224
Db 1501 GGGGCTCTGAACCCCAAGCCAGCCAGCCAGGAGCA-----GTGAGAGCCAG 1542
QY 225 LeuLeuMetProAspArgGlyLysGlyLysValLysIleGluSerTrpIleLysAspGlu 244
Db 1543 GCCCTGAAGACAAAAGGGATGAGAACCTGAGAGCATAGAGGGAAGTGAAGAACCAT 1602
QY 245 AlaGluThrAlaAlaValAlaTyrArgLysArgLeuMetThrGluGlyGlyLeuAla 264
Db 1602 ----- 1602
QY 265 AlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPheGlyValProSer 284
Db 1603 -----ATCTGTGAAGAAAAAGAGGCC 1623
QY 285 AsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySerAsnGluIleAla 304
Db 1624 GAGCTGAGCAGTTCAGCTCAGCAGCCCTCGGTTCATCCAGCAGCGTCCCAATATGTACATG 1683
QY 305 GlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIleValGluSerSer 324
Db ----- 324


```
QY 122 ValGluThrThrValThr-----ValSerGlnProSer 132
   |||
   :|
   :|
Db 465 ATGGAAGACTGCTTGGCCCATCTTGGAGAAAAGTGTCCAGGAACCTGAAGAGCCCTCTC 524
   |||
   :|
   :|
QY 133 GlnGluIleValProGluThrSerAsnLysPro-----GluGlyGlyArgMetCys 149
   |||
   :|
   :|
Db 525 CATAAAGCATTTGCAAAATCTCTGAGCCAGCGCATGATCATATCAGGCAATTCGGGAATGT 584
   |||
   :|
   :|
QY 150 GluLeu-----MetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsnIleSer 166
   |||
   :|
   :|
Db 585 ACATGGAGACACACAGTTTCATGCCAGCGCTGGAATAAGATTTGGTGCCTCTGGTTTG 644
   |||
   :|
   :|
QY 167 AspGlnAlaLysLeuMetGluIleProSerAlaLeuLysLeuAlaLysGluProAla 186
   |||
   :|
   :|
Db 645 CTACGACAAATGCTTTTGAA-----TTGACAGACTTGTCAAGAACCCTTG 692
   |||
   :|
   :|
QY 187 LysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGlyArgAla----- 203
   |||
   :|
   :|
Db 693 AGCGCACTGCTGCAG---TTTGGCGTGACATACCTGGAGGACTATTGGCAGAGTACATC 749
   |||
   :|
   :|
QY 204 -----PheThrLysGluSerPheLeuMetProAspArgGlyLysGlyLys 214
   |||
   :|
   :|
Db 750 ATTCAAGCAAGTGGCTGGGCACCTGTGTAGTTCTTGTAGTCAGAGGAGGAGGAATACCT 809
   |||
   :|
   :|
QY 215 GlnValSerLeuLeuLysSerPheLeuLeuMetProAspArgGlyLysGlyLys 234
   |||
   :|
   :|
Db 810 GGAATCCTGCAGAGATAGCAATGACATTTACATCTCTGCCCGGACCACTCTGGACAA 869
   |||
   :|
   :|
QY 235 ValLysIleGluSerTyrIleAspGluAlaGluThrAlaAlaValAlaTyrArgLys 254
   |||
   :|
   :|
Db 870 GTCAGT-----CCCCAGAGTCTCCAACTGTGACCACTTCTTCTGTCAGTCT 914
   |||
   :|
   :|
QY 255 ArgLeuMetThrGluGlyLeuAlaAlaGluLysMetAspAlaArgGlyLeuLeu 274
   |||
   :|
   :|
Db 915 GAGAGCTTACTGTG---TCACTGTGCGTACGACAGATTGGCACACAGAAAGCCTGCCA 971
   |||
   :|
   :|
QY 275 LeuLeuValAlaCysPheGlyValProSerAsnPheArgSerThrAspLeu---LeuAsp 293
   |||
   :|
   :|
Db 972 GTGTCTCATTAGC-----CCTGAGTCTCTGGCAGCAGATTGCAATGGATCCTGAA 1019
   |||
   :|
   :|
QY 294 LeuIleArgMetSerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeu 313
   |||
   :|
   :|
Db 1020 GAAGTGAAGAGCTTAGACAGCAAC-----GGAGCTGGAGAGAGAGAGTGAG--- 1064
   |||
   :|
   :|
QY 314 ValProMetValSerGlyIleValGluSerSerIleLysArgGlyMetHisIleGluAla 333
   |||
   :|
   :|
Db 1065 -----AACAACTCCTTAATCTGACATTTGTCAGCTGCAGTGAGAGAA 1103
   |||
   :|
   :|
QY 334 LeuGluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThr 353
   |||
   :|
   :|
Db 1104 GAAGAGGTG-----CCCGAGGGCATGGAAGAGCTGCTGGGCTTCTGTGGTCTTCCCA 1157
   |||
   :|
   :|
QY 354 SerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerPro 373
   |||
   :|
   :|
Db 1158 CGG-----CGGAGCTGCAAGNGGCATCTCTGAGCCCGAGCTCCC 1199
   |||
   :|
   :|
QY 374 LeuAlaPheLysGluAlaAlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCys 393
   |||
   :|
   :|
Db 1200 TTGCTTCCACATATCCTGCTCCACCTCCCTGCTGGGGACAGGGAACTGTGACACAGAGTG 1259
   |||
   :|
   :|
QY 394 MetGluThrHisLysLeuAspProAlaLys-----GluLeuProGlyTyrGln 409
   |||
   :|
   :|
Db 1260 ATCAGAGTTGAAATCCAGCCCTGCTACATCTCTGTTGTAGAACCTTGATGAAGAGAG 1319
   |||
   :|
   :|
QY 410 IleLys-----GluGlnIleValSerLeuGluLysAsp 420
   |||
   :|
   :|
Db 1320 GTGAAAGCAGCAACAACTGAACCTACTGAAGTGGAGGAGGTGTCCCGCACTGGAACCC 1379
   |||
   :|
   :|
QY 421 ThrLeuGlnLeuAspLysGluMetGluIleLysAlaArgSerLeuLeuMetGluGlu 440
   |||
   :|
   :|
Db 1380 ACAGAAACGCTGCTGAGTGAGAGAGAGATAACCCAGAGGAGAGAGCCCTTGTGGAGAG 1439
   |||
   :|
   :|
QY 441 AlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgPro----- 455
```

```
Db 1440 CTGTCCCTCGCCAGCGAG-----AAGAAGCCGTGCCCGCTCTGAG 1481
   |||
   :|
   :|
QY 456 -----ArgLeuSerProMet---GluMetProProValThrSerSer 468
   |||
   :|
   :|
Db 1482 GGCAAGTCTAGACTGTCCCCCGCGGTGAGATGAAGCCCATCGCTGTCT 1532
   |||
   :|
   :|
RESULT 12
US-09-098-901-1
; Sequence 1, Application US/09098901B
; Patent No. 6218144
; GENERAL INFORMATION:
; APPLICANT: Scott, Matthew
; APPLICANT: Sisson, John C.
; TITLE OF INVENTION: Costal2 Genes and their Uses
; FILE REFERENCE: SUN-65P
; CURRENT APPLICATION NUMBER: US/09/098,901B
; CURRENT FILING DATE: 1998-06-17
; EARLIER APPLICATION NUMBER: 60/051,347
; EARLIER FILING DATE: 1997-06-30
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3605
; TYPE: DNA
; ORGANISM: D. Melanogaster
US-09-098-901-1

Alignment Scores:
Pred. No.: 0.0205 Length: 3605
Score: 127.50 Matches: 150
Percent Similarity: 35.48% Conservative: 103
Best Local Similarity: 21.04% Mismatches: 244
Query Match: 4.11% Indels: 218
DB: 3 Gaps: 34

US-09-890-475-1 (1-609) x US-09-098-901-1 (1-3605)
QY 18 ProLeuLeuGlnArgHisGlnSerGluGlnArgArgGluLeuProLysIleValGlu 37
   |||
   :|
   :|
Db 1492 CCATTTCTGGAGCCGAGGAGCCCGAGGAGTGAATTCGAGAGCCCAATTCGGAGTCG 1551
   |||
   :|
   :|
QY 38 ThrGluSerThrSerMetAspIleThrIleGlnSerLysGlnProGlnPheLeuLys 57
   |||
   :|
   :|
Db 1552 CCAAACTCCGACAAACGAAACGACACAGCAATAGTCGATCGCCCGCATCTGGACGAC 1611
   |||
   :|
   :|
QY 58 SerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAsp 77
   |||
   :|
   :|
Db 1612 AGCATAGAAAGTCTA-----ATGGAAGAGTTTCGCCGACAAACAGAC 1653
   |||
   :|
   :|
QY 78 Asp-----LeuGlnLysHisIleGluSerIleGluAsn 88
   |||
   :|
   :|
Db 1654 GCTCTTATACTTGAAACACACGCTGAATATCTATCAAGCATCCGAAGCGGTATGCAA 1713
   |||
   :|
   :|
QY 89 AlaIleAspSerLysLeuGluSer-----AsnGly-----Val 99
   |||
   :|
   :|
Db 1714 AGCAAGACCGCGAGATTGAGGCGACAGCCGACAGAGAAATGGTGATGATCGAAAGTC 1773
   |||
   :|
   :|
QY 100 ValLeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerProProArgAsnAsn 119
   |||
   :|
   :|
Db 1774 AGCATTTGGCAGTCGACGAGGAAGTGT---CAGCCAGGTGTAGCTTAAGTACTGTCTGAG 1830
   |||
   :|
   :|
QY 120 ValSerValGluThrThrValThrValSerGlnProSerGlnGluIleValProGluThr 139
   |||
   :|
   :|
Db 1831 CTTGCCATGCTTAATCGGTAGCTTCCACAGACGCCCTCCGCCCATCATCGCTGAGTCG 1890
   |||
   :|
   :|
QY 140 -----SerAsnLysProGluGlyGlyArgMetCysGluLeuMetCys 153
   |||
   :|
   :|
Db 1891 GTCGTCGATCCTCTGAAAGAGTTCTTCGGCGGAGAGNATCCGTCAGCGGCTCTCGCTGCC 1950
   |||
   :|
   :|
QY 154 Ser-----LysGlyLeuArgLysTyrIleTyrAla 163
   |||
   :|
   :|
Db 1951 GCGCGCCCACTGCTCTCTATTGAACAGCTGCAGAAAAAATTCGCAAACTGGTCT---GCT 2007
```

QY 164 AsnIleSerAspGlnAlaLysLeuMetGluGluIle-----175
 Db 2008 GAGATCGAGGCAACACACGACAGTTCAGGGAATCGAAGAACAAATCCAGTAAACAAA 2067
 QY 176 -----ProSerAlaLeuLys-----Leu 181
 Db 2068 ATATAATCGCCGAATTG-GTCAAGAACACAGCGATACACGACGCCATGCAAGAACAAAGATTT 2126
 QY 182 AlalysGluProAlaLysPheValLeuAspCys-----IleGlyLys 195
 Db 2127 CACAGAAACGTCGCAAACTTGAGCGCCAGTGCACAAAGCCCAAGACGATGAGTTAG 2186
 QY 196 PheTyrLeuGlnGlyArg-----201
 Db 2187 CGCGTAGTTCAAGCCCGGGTCAGTCGAGATTGAGCGATGGACCACGATAATCGGACAT 2246
 QY 202 -----ArgAlaPheThrLysGluSerProMet-----SerSerAlaArg 214
 Db 2247 CTCGAGCGTCGACTAGAAAGACCTCAGCTCAATGAAGCATATTGCGGGTGAGAGCGGACAG 2306
 QY 215 GlnValSerLeuLeu-----IleLeuGluSerPheLeuLeuMetProAspArgGly 231
 Db 2307 AAGTGAAGAGAGTACAGCAATCGGTGGCGAGTCGCGAAACACAGCGCGATGATTACAG 2366
 QY 232 LysGlyLysValLysIleGluSerTrpIleLysAspGluAlaGluThrAlaAlaValAla 251
 Db 2367 AAA---AAGCTCGAAAGAGTGCAAGCTGCGCTGCGAGATGAGAGCGGAGCTGCTCAA 2423
 QY 252 TrpArgLysArgLeuMetThrGluGlyLeuAlaAlaGlu-----LysMet 268
 Db 2424 CTACGAGATCCAGGAGACTGGCAAGAGCTAGTGAAGCGCAAGGTTCTCCCGAGCAA 2483
 QY 269 AspAlaArgGlyLeuLeuLeuValAlaCysPheGlyValProSerAsnPheArgSer 288
 Db 2484 CAAGCCCGCGAGTTAAAGCGAGTACAGCT-----AGGATT 2519
 QY 289 ThrAspLeuLeuAspLeuIleArgMetSerGlySerAsn-----GluIleAlaGlyAlaLeu 307
 Db 2520 AGCCACTTAAATCATTACGAGAGAGTGGATACCTGGAGAGCAGCAGCGGACCA 2579
 QY 308 LysArgSerGlnPheLeu-----ValProMetValSerGlyIleValGluSerSer 324
 Db 2580 GAACAGCAGGACCTTCGTCATGATCGGCACTTGGCGGAACTCGTGACTTGTG 2639
 QY 325 IleLysArgGlyMetHis-----330
 Db 2640 TTGAAGAACGCTGTCATTGGACCGCAAACTTAAGCGGACAGGTCTGACGCCAAAG 2699
 QY 331 -----IleGluAlaLeuGluMetValTyrThr 339
 Db 2700 GAGGAGCGCAAGCTGCTCGAGTGGATGAGGCGCATCGAGGCGCATAGATCGGCCATAGAA 2759
 QY 340 PheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSer 359
 Db 2760 TTCAAGAACGATGATACACGGCGCCAGCGCTCCATCGACACAGGAC---CGAATTGAG 2816
 QY 360 LysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeuAlaPheLysGluAla 379
 Db 2817 CGGGAAGGGAGACAGATGCTGATGGCAGCGCTAAATCGTCTCTCAACGAGGAGATG 2876
 QY 380 AlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCys-----393
 Db 2877 CGAACACTTCTGTACAAATFACTTCAGAGGTTATCGATTTCGCGGACTTTCAGAAAG 2936
 QY 394 MetGluThrHisLysLeuAspProAlaLysGluLeuProGlyTrpGluIleLysGluGln 413
 Db 2937 CTGAGCTCAGCTGCTGTCAGTTGGCGCTGAGCGGATGCTCGGAGTGGAGGAGCGT 2996
 QY 414 IleValSerLeu-----GluLysAspThrLeuGlnLeu 424
 Db 2997 GTTCTGTCAATCGCGTGGCCAGGCTAGACTGGAAGCGCAACGGAATCGCGTGTGCTG 3056

QY 425 AspLysGluMetGluGluLysAlaArgSerLeuSerLeuMet-----Glu 439
 Db 3057 CAGCGCCAGCAGCAATGAAA-----CTCATTGTGATGCTGCTCAGATGCGGAG 3107
 QY 440 GluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSerPro 459
 Db 3108 GAAAGCTGGCCAGTTCGGCCAGCTACGAGAACGAGCTTTGGCCCT-----GCC 3158
 QY 460 MetGluMetProValThrSerSerSer-----TyrSerProIleLysArgAsp 476
 Db 3159 TGTGTGCGCCCGCGCTGAGCTAGTATTTCGACTACGATCATTTCTACAAAGGT 3218
 QY 477 ArgSerPheProSerGluArgAspAspGlnAspGluIleSerAlaLeuValSerSer 496
 Db 3219 GCGGCAATCCAGCAAG-----GCATGATC-----3245
 QY 497 TyrLeuGlyProSerThrSerPheProHisArgSerArgArgSerProGluTyrMetVal 516
 Db 3246 -----AAAGCCCAAG-----3257
 QY 517 ProLeuProHisGlyGlyLeu-----GlyArg 525
 Db 3258 CGATGCCCGCCGCTCGGCGCTAGACAAATACAGGACAAAGCAACGACGCGGAGCC 3317
 QY 526 SerValTyrAlaTyrGluHisLeu-----AlaProAsnSerTyrSer 539
 Db 3318 AACATCTTTGCCAAGTTCATGCTCACCAGATATGCTGAGTCCGCGCAGCGGTTC 3377
 QY 540 ProGly-HisGlyHisArgLeuHisArgGlnTyrSerProSerLeuVal-----HisG 557
 Db 3378 TCAGGTCACCGCGCGAGGATCCAC-----GGCCCTGATTGAGTCAACCCACG 3428
 QY 557 LysGlnArgHisProLeuGlnTyrSerProProIle 568
 Db 3429 GCCA-CGGCAACCACTACGTCGACCAACCACTG 3462

RESULT 13

US-09-919-172-97
 ; Sequence 97, Application US/09919172
 ; Patent No. 6673545
 ; GENERAL INFORMATION:
 ; APPLICANT: Paris, Mary
 ; APPLICANT: Turner, Christopher M.
 ; TITLE OF INVENTION: PROSTATE CANCER MARKERS
 ; FILE REFERENCE: PA-0036 US
 ; CURRENT APPLICATION NUMBER: US/09/919.172
 ; CURRENT FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: 60/222.469
 ; PRIOR FILING DATE: 2000-07-28
 ; NUMBER OF SEQ ID NOS: 102
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 97
 ; LENGTH: 10432
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. 6673545 2700132CB1
 US-09-919-172-97

Alignment Scores:
 Pred. No.: 0.121 Length: 10432
 Score: 127.50 Matches: 138
 Percent Similarity: 33.43% Conservative: 91
 Best Local Similarity: 20.15% Mismatches: 230
 Query Match: 4.11% Indels: 227
 DB: 4 Gaps: 30

US-09-890-475-1 (1-609) x US-09-919-172-97 (1-10432)

QY 11 GlnProThrThrThrAlaAsnProLeuLeuGlnArgHisGlnSerGluGlnArgArg---29
 Db 2717 GAGCCTTCAAAAAACATATCCACTGTAAACAGGTAGGAGGTCTACAGATTTCAGGAAT 2776


```
/ CURRENT APPLICATION NUMBER: US/09/976,594
/ CURRENT FILING DATE: 2001-10-12
/ PRIOR APPLICATION NUMBER: 60/240,409
/ PRIOR FILING DATE: 2000-10-12
/ NUMBER OF SEQ ID NOS: 1143
/ SOFTWARE: PERL Program
/ SEQ ID NO 21
/ LENGTH: 10432
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Inbyte ID No. 6673549 2700132CB1
US-09-976-594-21

Alignment Scores:
Score: 0.121 Length: 10432
Pred. No.: 127.50 Matches: 138
Percent Similarity: 33.43% Conservative: 91
Best Local Similarity: 20.15% Mismatches: 230
Query Match: 4.11% Indels: 227
DB: 30 Gaps: 30

US-09-890-475-1 (1-609) x US-09-976-594-21 (1-10432)
QY 11 GlnProThrThrThAlaAsnProLeuLeuGlnArgHisGlnSerGluGlnArgArg--- 29
Db 2717 GAGCCTTCAAAACAGTATCCACTGTAACAGGTGAGGAGGTCTACAGAGTTCAGGAAT 2776
QY 30 ---ArgGluLeuProLysIleValGluThrGluSerThrSerMetAspIle----- 45
Db 2777 ATACAGAGCTACCTGTGGAAGTAGTAGAGTGAAGAAACAAATACAGAAATGTTGAGTGC 2836
QY 46 -----ThrIleGlyGlnSerLysGlnProGlnPheLeuLys 57
Db 2837 ATCTTAAAGAGGTGAGAGGACACACTACTACCAAGAGAGAGAGGAGATGAAG 2896
QY 58 SerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAsp 77
Db 2897 GAAATAGAA-----AGACCTTTTGAG 2917
QY 78 AspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsn 97
Db 2918 ACATATAAGGAAATATTGAAATTAAGAGAAAC-----GATGAAAGATGAAGCA--- 2968
QY 98 GlyValValLeuAlaAlaAsnAsnAsnAsnPheHisGln-----ProMetLeuSer 114
Db 2969 -----ATGAGAGATCAAGAACTTGGGGCGCAAAATGTGCCCAATG----- 3010
QY 115 ProProArgAsnAsnValSerValGluThrThrValThrValSerGlnProSerGlnGlu 134
Db 3011 -----TCTGACCTGACAGACCTCAAGAGCTTGCTGATACAGAA 3049
QY 135 IleValProGluThrSerAsn----- 141
Db 3050 CTCATGAAGACGCGCAGCTGGCGCAAGATCTCTCCAAACCCCAAGATCATGCCAAGCA 3109
QY 142 ---LysProGluGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyr 160
Db 3110 CCAAGAGTGAGAAAGGCAAAATCACTAAATGCTCCAGTCCAGTCAATCAACCAAGCA 3169
QY 161 IleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAlaLeuLys 180
Db 3170 ATAACACCCCAACACACACAAACACAGTTG-----AAGCATCC 3211
QY 181 LeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPhe-----TyrLeuGln 199
Db 3212 CTGGGAAAGTAGGTGTGAAGAGAGCTCTAGCAGTCCGCAAGTTCACACGGAGCTCA 3271
QY 200 GlyArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeu 219
Db 3272 GGGGAGACCACGACACGACAGAGAGCCAGCAGGAGATGGCAAGAGCATC----- 3322

/ CURRENT APPLICATION NUMBER: US/09/976,594
/ CURRENT FILING DATE: 2001-10-12
/ PRIOR APPLICATION NUMBER: 60/240,409
/ PRIOR FILING DATE: 2000-10-12
/ NUMBER OF SEQ ID NOS: 1143
/ SOFTWARE: PERL Program
/ SEQ ID NO 21
/ LENGTH: 10432
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Inbyte ID No. 6673549 2700132CB1
US-09-976-594-21

Alignment Scores:
Score: 0.121 Length: 10432
Pred. No.: 127.50 Matches: 138
Percent Similarity: 33.43% Conservative: 91
Best Local Similarity: 20.15% Mismatches: 230
Query Match: 4.11% Indels: 227
DB: 30 Gaps: 30

US-09-890-475-1 (1-609) x US-09-976-594-21 (1-10432)
QY 11 GlnProThrThrThAlaAsnProLeuLeuGlnArgHisGlnSerGluGlnArgArg--- 29
Db 2717 GAGCCTTCAAAACAGTATCCACTGTAACAGGTGAGGAGGTCTACAGAGTTCAGGAAT 2776
QY 30 ---ArgGluLeuProLysIleValGluThrGluSerThrSerMetAspIle----- 45
Db 2777 ATACAGAGCTACCTGTGGAAGTAGTAGAGTGAAGAAACAAATACAGAAATGTTGAGTGC 2836
QY 46 -----ThrIleGlyGlnSerLysGlnProGlnPheLeuLys 57
Db 2837 ATCTTAAAGAGGTGAGAGGACACACTACTACCAAGAGAGAGAGGAGATGAAG 2896
QY 58 SerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAsp 77
Db 2897 GAAATAGAA-----AGACCTTTTGAG 2917
QY 78 AspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsn 97
Db 2918 ACATATAAGGAAATATTGAAATTAAGAGAAAC-----GATGAAAGATGAAGCA--- 2968
QY 98 GlyValValLeuAlaAlaAsnAsnAsnAsnPheHisGln-----ProMetLeuSer 114
Db 2969 -----ATGAGAGATCAAGAACTTGGGGCGCAAAATGTGCCCAATG----- 3010
QY 115 ProProArgAsnAsnValSerValGluThrThrValThrValSerGlnProSerGlnGlu 134
Db 3011 -----TCTGACCTGACAGACCTCAAGAGCTTGCTGATACAGAA 3049
QY 135 IleValProGluThrSerAsn----- 141
Db 3050 CTCATGAAGACGCGCAGCTGGCGCAAGATCTCTCCAAACCCCAAGATCATGCCAAGCA 3109
QY 142 ---LysProGluGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyr 160
Db 3110 CCAAGAGTGAGAAAGGCAAAATCACTAAATGCTCCAGTCCAGTCAATCAACCAAGCA 3169
QY 161 IleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAlaLeuLys 180
Db 3170 ATAACACCCCAACACACACAAACACAGTTG-----AAGCATCC 3211
QY 181 LeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPhe-----TyrLeuGln 199
Db 3212 CTGGGAAAGTAGGTGTGAAGAGAGCTCTAGCAGTCCGCAAGTTCACACGGAGCTCA 3271
QY 200 GlyArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeu 219
Db 3272 GGGGAGACCACGACACGACAGAGAGCCAGCAGGAGATGGCAAGAGCATC----- 3322

220 IleLeuGluSerPheLeuLeuMetProAsp-----ArgGlyLys 232
3323 -----AGAACGTTTAAAGAGTCTCCAAAGACAGATCCTGGACCAGCAGCCCGTGTAACT 3376
QY 233 GlyLysValLysIleGluSerTPIleLysAspGluAlaGluThrAlaAla---ValAla 251
3377 GGAATGAGAGAGTGGCCAGACGCTTAAAGAGAGGCCAGTCACCTAGAACCTGGGT 3436
QY 252 TrpArgLysArgLeuMetThrGluGlyLeuAlaAlaLysMetAspAlaArg 271
3437 GGGTTCAAAGAGCTCTTCCAGACACAGGTCCTCTGAGGAATCAATGACTGATGAGAA 3496
QY 272 GlyLeuLeuLeuValAlaCysPheGlyValPro----- 283
3497 -----ACTACCAAAATAGCTGCAAAATCTCCACCACAGATCACTGACACTCCAA 3550
QY 284 -----SerAsnPheArgSerThrAspLeu-----Leu 292
3551 AGCACAAAGCAATGGCTTAAGAGAGTCTCAGGAAAGACAGATGATAGGAGAGAAATCTT 3610
QY 293 AspLeuIleArgMetSerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPhe 312
3611 GCACCTCAGGAACTAAACACCATCA-----GCAGGG-----AAAGCCATG 3649
QY 313 LeuValProMetValSerGlyIleValGluSerSerIleLys-----ArgGlyMetHis 330
3650 CTTACGCCCAACACAGCAGGAGGTGATGAGAAAGACATTAAGCATTTATGGGAATCCA 3709
QY 331 IleGluAlaLeuGluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAlaLeu 350
3710 GTGCAGAACTGGACCTGGCAGGAACCT----- 3736
QY 351 ValLeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLysAla 370
3737 -----TTACCTGCAGCAAAAGACAGACTACAGACTCTCCAGACTCTCGTAAAGAAAGGCC 3781
QY 371 GlnSer-----ProLeuAla-----PheLysGlu----- 378
3782 CAGGCTCTAGAACAGCTGGTGGCTTTAAAGAGCTCTCCAGACTCTCGTAAAGAAAGGCC 3841
QY 378 ----- 378
3842 GAATTAGTGGCTGCTGTAAACACCTAAATACCTCGACTCTCCAGACTCTCCAGACTCAG 3901
QY 379 -----AlaAlaThrLysGln----- 383
3902 GTGGACACCCCAACAGCACAAAGACAGCCCAAGAGAGTATCAGGAAAGCAGATGTA 3961
QY 384 -----LeuAla-----ValLeuSerSerValMetGlnCysMetGluThr 396
3962 GAGGAGAACTCTTAGCTGCAGGAATCTAATGCCATCAGCAGGCAAGCCATGCAACG 4021
QY 397 -----His 397
4022 CCTAAACCATCAGTAGGTGAGAGAGAGACATCATATTTTGGGAACTCCAGTGCAG 4081
QY 398 LysLeuAspProAlaLysGluLeuProGlyTrpGlnIleLysGluGlnIleValSerLeu 417
4082 AAACCTGGACCTGCAGAGAACTTAAACCGCAGCAGAGAGCGGCCCAACAACTCTTAAG 4141
QY 418 GlyLysAspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeu 437
4142 GAGGCCAGGCTCTGGAAGACCTGAGCTGGCTTTTAAAGAGCTCTTCCAGACCCCTGGTCA 4201
QY 438 MetGluGluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLeu 457
4202 ACTGAGAGACAGCTGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4240
QY 458 SerProMetGluMetProProValThrSerSerSerTyrSerProIleTyrArgAspArg 477
4241 CCCTGCAATCTTCTCCACCAAGATCAGCAGACACCCCAACAGCAGCAAGAGGAGCC 4300
QY 478 SerPheProSerGlnArgAspAspGlnAspGluIleSerAlaLeuValSerSery 497
```

Db 4301 AAGACACCTTTGGAGAAAGGACGTACAGAGGAGCTCTCAGCCCTGAAGAGCTCA-- 4358
Qy 497 rLeuGlyProSerThrSerPheProHisArgSerArgSerProGluTyrMetValPr 517
Db 4359 -----CACACATCATCAGGGGAAACACACACAGATATAA 4393
Qy 517 oLeuProHisGlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAlaProAsnSe 537
Db 4394 -GTACCA---GGAGGTGAGGATAAAGCATCAACGGCTTTAGGGAACTGCAAAACAGAA 4449
Qy 537 rTyrSerProGlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGl 557
Db 4450 ACTGGACCCA-----GCAGCAAGTGTACTGTGTAG 4479
Qy 557 yGlnArgHisPro 561
Db 4480 CAAGAGGCACCCA 4492

RESULT 15
US-08-466-390-3
; Sequence 3, Application US/08466390
; Patent No. 5686562
; GENERAL INFORMATION:
; APPLICANT: TOKUATLY, GARY
; APPLICANT: LIGARD, GRAHAM P
; TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
; TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: 125 HIGH STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,390
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER ESQ, EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: MTP-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6306 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..6306
; PUBLICATION INFORMATION:
; AUTHORS: COMPTON, DUANE A
; AUTHORS: SZILAK, ILYA
; AUTHORS: CLEVELAND, DON W
; TITLE: PRIMARY STRUCTURE OF NUMA, AN INTRANUCLEAR
; TITLE: PROTEIN THAT DEFINES A NOVEL PATHWAY FOR
; TITLE: SEGREGATION OF PROTEINS AT MITOSIS
; JOURNAL: J. Cell Biol.
; VOLUME: 116
; PAGES: 1395-1408
; DATE: 1992

US-08-466-390-3
Alignment Scores:
Pred. No.: 0.0587 Length: 6306
Score: 127.00 Matches: 161
Percent Similarity: 34.61% Conservatives: 102
Best Local Similarity: 21.18% Mismatches: 275
Query Match: 4.09% Indels: 222
DB: 1 Gaps: 35
US-09-890-475-1 (1-609) x US-08-466-390-3 (1-6306)
Qy 9 AlaAlaGlnProThrThrThrAlaAsnProLeuGlnArgHisGlnSerGlu----- 26
Db 3331 GCTGCTGCAGCAGACAGAGCCCAACAGGCCCAAGCTTGAAGCAGCTGCGGCGCAGAGGTGAGC 3390
Qy 27 -----GlnArgArgGlu-----LeuProLysIleVal 36
Db 3391 AAGCTGGAACAGCAATGCCAGAGCAGCAGGAGGCTGACAGCTTGGACGCGCTC 3450
Qy 37 GluThrGluSerThrSer-----MetAspIleThrIleGlyGln 49
Db 3451 GAGGCTGAGCGGGCTCTCCGGGCTGAGCGGACAGTGTCTTGGAGACTCTCGAGGGCCAG 3510
Qy 50 -----SerLysGlnProGlnPheLeuLysSerIleAsp 60
Db 3511 TTAGAGGAGAAGGCCAGGAGCTAGGGCACAGTCAGAGTGCCTTAGCTCGGCCCAACGG 3570
Qy 61 GluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAspAspLeuGln 80
Db 3571 GAGTTGGCTGCTTCCGCCAACCAAGTACAGACCACAGCAGAGGTGAAGATGAGTGGAG 3630
Qy 81 LysHisIle-----GluSerIleGluAsnAlaIleAspSerLysIle 94
Db 3631 GCCCAGGTGGCGCGCGCGCAGAGCTGAGAGGAAATAGCTCATCAGCAGCTTG 3690
Qy 95 GluSerAsnGlyValValLeuAlaAlaArgAsnAsnPheHisGlnProMetLeuSer 114
Db 3691 GAGGAGGAGGTGTCATCTCTG-----AATCGCAGGTCTCTGGAG 3729
Qy 115 ProProArgAsnAsnValSerValGluThrThrValThrVal-----SerGlnProSerGln 133
Db 3730 AAGGAGGGGAGAGCAGAGAGTGAAGCGGCTGTGATGCGCGAGTCAAGAACGCGCAG 3789
Qy 134 GluIle-----ValProGluThr 139
Db 3790 AAGCTGGAGGAGAGCTGCGCTGTCAGGCGAGAGCAGCAGCAGTCCAGAGCTG 3849
Qy 140 SerAsnLysPro---GluGlyGlyArgMetCys----- 149
Db 3850 CAGAACGCGAGCTCTGCTCTGCGGGAGGAGGTGCGAGCGCTCCGGAGGAGGCTGAGAAA 3909
Qy 150 GluLeuMetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAla 169
Db 3910 CAGCGGTGGCTTCAGAACTCTGCGCAG-----GAGCTGACCTCAGAGCT 3957
Qy 170 LysLeuMetGluGluIleProSerAlaLeuLysLeuAlaLysGlu----- 184
Db 3958 GAGCGTGGGAGGAGCTGGCGCAAGATTGAAGGCGTGGCAGGAGAGTCTTCCAGAAA 4017
Qy 184 ----- 184
Db 4018 GAGCAGGCCCTCTCCACCTCGAGCTCGACACACACAGCAGCAGCGCCCTGGTGTAGT 4077
Qy 185 -----ProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGlyArg 202
Db 4078 CTGCTGCCAGCTAAGCACCTCTGCGCAG-----CAGCTGACGCGCAGCAG 4122
Qy 203 AlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuIle----- 220
Db 4123 GCCCGTGGCGAGAAACGCCACCGTGGAGGAGTGGAGCAGACAGCAAGCAGCGCGTGGGGA 4182
Qy 221 LeuGluSerPheLeuLeuMetProAspArgGlyLysGlyLys---ValLysIleGluSer 239

Pred. No.: 1,43e-74 Length: 2368
Score: 773.50 Matches: 162
Percent Similarity: 64.38% Conservativity: 91
Best Local Similarity: 41.22% Mismatches: 108
Query Match: 24.92% Indels: 32
DB: 12 Gaps: 5

US-09-890-475-1 (1-609) x US-10-424-599-59479 (1-2368)

Qy 57 LysSerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPhe 76
Db 195 AAATCGGTGACAACTCAACAGCCTCGCCTTCGATACAGCCTTCAATACAGGTAC 254
Qy 77 AspAspLeuGlnLysHisIleGluSerIleGluAlaIleAspSerLysLeuGlu--- 95
Db 255 GAAGAAATTCAGAAACACCTCGAATTCATCGAACAGCCATTCACAGGAGACCAAGGAG 314

Qy 96 -----SerAenGlyValValLeuAla 102
Db 315 CTCGAGCAGTACGCTACAAATCTACTCAAGAAACCGCGAAACCGCGCTCGTTCAATCA 374

Qy 103 AlaArgAsnAenAsnPheHisGlnProMetLeuSerProArgAsnAenValSerVal 122
Db 375 -----GATTCAAAATCCGAGGCCA 392

Qy 123 GluThrThrValThrValSerGlnProSerGlnGluIleValProGluThrSerAsnLys 142
Db 393 GAAGAAGCCAAAGCGGTGAGAGAAAGAGAAAGGTAAAGAAAGAAAGAAAGAAAGAA 452

Qy 143 ---ProGluGlyGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTrpIle 161
Db 453 GAAGACGAGCTTATACGCTTTGCAAAACGATGATAGCCGAGCGCTCGTAAATACGTG 512

Qy 162 TyrAlaAenIleSerAspGluAlaLysLeuMetGluGluIleProSerAlaLeuLysLeu 181
Db 513 TTAACGCGCTATCCGAAACCGCGCTTCGGGAACAGGTACCCGTTGCGCTGAGGAGC 572

Qy 182 AlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGlyArg 201
Db 573 CGCGGAGAGCCTCAAGCTGCTGTTGAATGCAATGGAGGTTTTCCTTCAGGGAGC 632

Qy 202 ArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuLeu 221
Db 633 AAAGCTTACAGAAAGAACTCCCGATGTTCCCGCAAGCAGAGTTTCGGTCTGTTTG 692

Qy 222 GluSerPheLeuMetProAspArgGlyLysValLysIleGluSerTrpIle 241
Db 693 GAGCACTACTTCTCTGCTGCTGGTGGGAATGAGAG---GACGTGGAGGCTTCATTG 749

Qy 242 LysAspGluAlaGluThrAlaAlaValAlaTrpArgLysArgLeuMetThrGluGlyGly 261
Db 750 AAGAGGAGGCGGATTCGCGCGGCTTCGCTGAGGAGAGAGATGTTTGTGAAGGCGGT 809

Qy 262 LeuAlaAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPheGly 281
Db 810 TTGCTGAAGGAGCAGTGAAGTGTGACCGAGGGTTGATTTCTTCCTCGCGCGCTTCGGG 869

Qy 282 ValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySerAsn 301
Db 870 ATTCACAGTGTTCACAGATGAGATATATACAACTTGTTGTTGTCAGCAATGGCAGA 929

Qy 302 GluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIleVal 321
Db 930 GAATTCCTCGATGCGCTCTCCAGTCTCAGCCCTCGCTTAAGAGGGTTTCAGATGTTGCA 989

Qy 322 GluSerSerIleLysArgGlyMetHisIleGluAlaLeuGluMetValTyrThrPheGly 341
Db 990 GATGGGATGACAAAGAGGATGCCCTTAAAGCTGTGATTTGGCTTATACCTTTGGG 1049

Qy 342 MetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLysGlu 361
Db 1050 TTTGAAGAGAAATATCTTCTCAGACAGCTCTGATTCATTCTCTGAGAGTCTGAAGAA 1109

Qy 362 SerPheGluArgAlaLysArgLysAlaGlnSer---ProLeuAlaPheLysGluAlaAla 380
Db 1110 ACTTGGAAAGAACCAACAGACGCGACGTGATTTTCCTAGTCACTGAGGTAGCACAT 1169

Qy 381 ThrLysGlnLeuAlaValLeuSerValMetGlnCysMetGluThrHisLysLeuAsp 400
Db 1170 GAAATAATTTGGCTGCTTTCATCTCTAGTAAATGTTTGGAGGTCACAGATTGAC 1229

Qy 401 ProAlaLysGluLeuProGlyTrpGlnIleLysGluGlnIle-ValSerLeuGluLysAs 420
Db 1230 CCTGTAAATTTCTGCTGCTGGTGGCACTTAAGATAAGATTACTCAACTTGGAGAAAG 1289

Qy 420 pThrLeuGlnLeuAspLysGluMetGluGluLysAla 432
Db 1290 TATTAATGATGCCAATAAAAAAATTTGATGAGAGTCA 1326

RESULT 2
US-10-425-114-21596
Sequence 21596, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313) B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 21596
LENGTH: 2041
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3354-037-B3_FLI
US-10-425-114-21596

Alignment Scores:
Pred. No.: 1,63e-30 Length: 2041
Score: 373.00 Matches: 157
Percent Similarity: 40.06% Conservativity: 103
Best Local Similarity: 24.13% Mismatches: 231
Query Match: 12.02% Indels: 158
DB: 12 Gaps: 27

US-09-890-475-1 (1-609) x US-10-425-114-21596 (1-2041)

Qy 5 ProProThrValAlaAlaGlnProThrThrThr-----AlaAsnPro 18
Db 123 CCTCTCTCTTCAGCAGCAGCGCGGCGCACACCTCGATATCTTCCCTCGGACCG 182

Qy 19 LeuLeuGlnArgHisGlnSerGluGlnArgArgGluLeuPro----- 33
Db 183 -----GAGCGAGCAGCAGCCTCCAGCGCTCCGCCCTCCCGCTCGTTATTGGGA 236

Qy 34 -----LysIleValGluThrGluSerThrSerMetAspIleThrIleGlnSerLys 51
Db 237 GGATTCGCATCTCTGACATGAGTCCGTCGCGAGCTCTTATGAACTCGACAGCTCCAG 296

Qy 52 GlnProGlnPheLeuLysSerIleAspGluLeuAlaAla----- 64
Db 297 ATACAGCAGCTTCAGGAGGCAATTTGCTGAGCTTGAGAGCCAGAGCAGCGGTTTCCATGAC 356

Qy 65 -----PheSerValAlaValGluThrPheLysArgGln 75
Db 357 CTCAAAGTGAAGCAGCTTGAAGACCACTTTCTGTGCTTGAACAGTCCCTCAAGAAAG 416

Qy 76 PheAspAspLeuGln----- 80
Db 76 PheAspAspLeuGln----- 80

Qy	13	ThrThrAlaAsnProLeuLeuGlnArgHisGlnSerGluGlnArgArgGlu	--- 31
Db	405	ACAAGGAAGGCTCAGGAGATACCTGGAGAAGCGCAAGCAGCTCTTTATGTCGAAGAGCAA	464
Qy	32	-----LeuProLysIleValGluThrGluSerThrSerMet	--- AspIleThrIleGly 48
Db	465	GCACGCTGGCAGAGGCTTCAGAGAAAGGAGTGTCTGTATTCACATTTGGAATGCT	524
Qy	49	GlnSerLysGlnProGlnPheLeuLysSerIleAspGluLeuAlaPheSerValAla	68
Db	525	CGAGAAAGCAGAGGAGGTC-----ACGATAAGTGATTGGCTATTGCTCCCAATGGG	578
Qy	69	ValGlu---ThrPheLysArgGlnPheAspAspLeuGlnLysHisIleGluSerIleGlu	87
Db	579	GGTAAGGGGACATTT-----CATGTGGAG-----GAC	605
Qy	88	AsnAlaIleAspSerLysLeuGluSerAsnGlyValValLeuAlaAlaArgAsnAsnAsn	107
Db	606	AAACCAGTGATGCT-----GTGTCTTTCAGCTAATGCTAAT	644
Qy	108	PheHisGlnProMetLeuSerProProArgAsnAsnValSerValGluThrThrValThr	127
Db	645	GTGGAAGAGGTGGTACTTTCTCTGAAAATCGAAATGTGAGTTG-----	689
Qy	128	ValSerGlnProSerGlnGluIleValProGluThrSerAsnLysProGluGlyGlyArg	147
Db	690	-----TCTTATCCAGACTTGGTAAAA	710
Qy	148	MetCysGluLeuMetCysSerLysGlyLeuArgLysTyrlleTyrlaAlaSerAsp	167
Db	711	CTTTGTAAAGAGATGGATGCTGTGGGCTTCACAAATTCATATCTGCACACCGTAAGAAC	770
Qy	168	GlnAlaLysLeuMetGluGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLys	187
Db	771	CTTGCTGCTGTGAAGGAGGAAATACCAATGCATTAAGACAGCTCTCTAACGGCGCTGT	830
Qy	188	PheValLeuAspCysIleGlyLysPheTyrlleLeuGln-----GlyArgArgAlaPheThr	205
Db	831	TTAGTTTTAGATCTCTGAAGGGGTTTTACTGTCACAGAAGTGTCAAATCAGGACGTAAAG	890
Qy	206	LysGluSerProMetSerSerAlaArgGlnValSerLeuLeuIleLeuGluSerPheLeu	225
Db	891	AAGATGCTAACTTACTGGGTGTTCGCGCAACATGTATCATGTGTGATGGAAATGCTATGT	950
Qy	226	---LeuMetProAspArgGlyLysGlyLysValLysIleGluSerTrpIleLysAspGlu	244
Db	951	GATTTCTCGAGCACTCAGGTGTGTGTCTTAAGTAATTTTCAGAGNATATCAGAGACCGG	1010
Qy	245	AlaGluThrAlaAlaValAlaAlaTrpArgLysArgLeu---MetThrGluGlyGlyLeuAla	263
Db	1011	GCTAAGCGAGTGTGTGAAGATGAAACCCAGATTTGATGCTCTTCACATCGATGTAGC	1070
Qy	264	AlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuLeuValAlaLysPheGlyValPro	283
Db	1071	AATGTGAATTCCTTGGAGGCTCATGCAATTTTGCACATCTTAGCCAGTTTGTGATTGCC	1130
Qy	284	SerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySerAsnGluIle	303
Db	1131	TCTGGTTTTAATGAGGAGGATTTATCTAGGCTGATTCCAATGGTATCTCGACGTCCCAA	1190
Qy	304	AlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIleValGluSer	323
Db	1191	ACTGCTGATTTATGCTGTGTGGGTGTGCAAAAAGATGCTGGTGTCATTTGAAGTT	1250
Qy	324	SerIleLysArgGlyMetHisIleGluAlaLeuGluMetValTyrllePheGlyMetGlu	343
Db	1251	TTGCTGTAATAGTGGCGGCAAAATGATGCTCTTAACTTGGCTTTTGCATTTGATCTTACA	1310
Qy	344	AspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLysGluSerPhe	363
Db	1311	GACCAATTTTCCACCCTTCTCTTACTGAAGTCTTACTTGAAGGATGCTGACAAAATTTCT	1370

Qy	364	GluaArgAlaLysArgLysAlaGlnSerProLeuAlaPheLysGluAlaIaThrLysGln	393
Db	1371	TCCTCCTGTCAGAAAGTGTTAACTCATCTCCCACTGCACAGATTGATGTTAATGATCGAGAG	1430
Qy	384	LeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeuAspProAlaLys	403
Db	1431	CTGATTGCACCTTAAAGCTGTAACTCAAGTGCATTGAAGATCATAACTGAT	1484
Qy	404	GlulLeuProGlyTrpGlnIleLysGluGlnIleValSerLeuGluLysAspThrLeuGln	423
Db	1485	CAGTATCCT	1523
Qy	424	LeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGluGluAlaAlaLeu	443
Db	1524	CTAGAGAGGCCAAAGCTGCAAGAGAGGGAA	1568
Qy	444	AlaLysArgMetTyrAsnGlnGlnIleLysArgProGluSerProMetGluMetPro	463
Db	1569	-----AAGCCTCAACCCAGAGACCCCGTCCAAATGGTGTG-----	1604
Qy	464	ProValThrSerSerSerTyrSerProIleTyrArgAspArgSerPheProSerGlnArg	483
Db	1605	-----GGATATGTCCA-----	1616
Qy	484	AspAspGlnAspGluIleSerAlaLeuValSerSerTyrLeuGlyProSerThrSer	503
Db	1617	-----CGTGTCACTAACATTCTTCGGACAAAACCTTCC	1649
Qy	504	PheProHisArgSerArgSerProGluTyrMet-ValProLeuProHisGlyGlyLe	523
Db	1650	TATGCTAGAGTTGTGCAGGTATCTCTCAATACGTGTATGATCAGCAGCCCT	1698
Qy	523	uGlyArgSerValTyrAlaTyr-----GluHisLeuAlaPro-----	535
Db	1699	-----TACATGTACCTGCACCAACTGAGAATCATTGCCCCCTCTCATCAG	1745
Qy	536	-----AsnSerTyrSerProGlyHisGlyHisArgLeuHisArg-----GlnTyr	550
Db	1746	CACCTGCAACATATAACATCTCTCCAGCCATGCGCAACTACTTCTGGAAATGGGTATCAGTA	1805
Qy	550	rSerProSerLeuValHis	556
Db	1806	CCAAAGCCTCATATCTCCAC	1824
RESULT 4			
US-10-425-114-30101			
; Sequence 30101, Application US/10425114			
; Publication No. US20040034888A1			
; GENERAL INFORMATION:			
; APPLICANT: Liu, Jingdong			
; APPLICANT: Zhou, Yihua			
; APPLICANT: Kovalic, David K.			
; APPLICANT: Screen, Steven E			
; APPLICANT: Tabaska, Jack E			
; APPLICANT: Cao, Yongwei			
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With			
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement			
; FILE REFERENCE: 38-21(53313)B			
; CURRENT APPLICATION NUMBER: US/10/425,114			
; CURRENT FILING DATE: 2003-04-28			
; NUMBER OF SEQ ID NOS: 73128			
; SEQ ID NO 30101			
; LENGTH: 2089			
; TYPE: DNA			
; ORGANISM: Glycine max			
; FEATURE:			
; OTHER INFORMATION: Clone ID: UC-GMROFIC109C08_FLI			
US-10-425-114-30101			
Alignment Scores:			
Pred. No.:		3,56e-29	2089
Score:		361.00	137
Percent Similarity:		39.39%	97

Best Local Similarity: 23.06% Mismatches: 236
 Query Match: 11.63% Indels: 124
 DB: 12 Gaps: 17

US-09-890-475-1 (1-609) x US-10-425-114-30101 (1-2089)

Qy	37	GluThrGluSerThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeu	56	311	GlnPheLeuValProMetValSerGlyIleValGluSerSerIleLysArgGlyMetHis	330
Db	253	GACTCGGATTCAGTTCACCTGACCTGAGATTCCTACTCTTCAAGATCAACACGCTGCGAG	312	1270	CTTGGGCTGTCAGAAAAGATCGCTGGTGTAAATTGAGGTTTTTGGTGAACAGCTGGCGCCACAA	1329
Qy	57	LysSerIleAspGluLeuAlaPheSerValAlaVal-----	69	331	IleGluAlaLeuGluMetValThrPheGlyMetGluAspLysPheSerAlaAlaLeu	350
Db	313	AAGGCATTGCTGAACCTTGAAGTTATCGGGCTGTACTCTTAACTTGAATGGAAGAA	372	1330	ATTGATGCTGTAACTTGGCTTTTGCATTTGATCTTACTGAACAGTTTTTCTCTCTCTTCT	1389
Qy	70	-----GluThrPheLysArgGlnPheAspAspLeuGln	80	351	ValLeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLysAla	370
Db	373	CTAGAGAAACATTTCCATCGTCTTGAAGAAATCCTTGAAGAGCGCTTTGATGAATTGAA	432	1390	TTACTGAAGTCTTACTTGAAGATCTAGAAAAGCTTCTTCTCTCTTGTAGAGTGTCTCAC	1449
Qy	81	LysHisIleGluSerIleGluAsnAlaIleAspSerLys-----LeuGluSerAsn	97	371	GlnSerProLeuAlaPheLysGluAlaAlaThrLysGlnLeuAlaValLeuSerSerVal	390
Db	433	GACCAAGAGAAGAGTTTGAAACAAACAGCAGGAGGCTCGTGAGATCTCGAAGCGG	492	1450	TCTTCTCCACTGCACAGATTGAGTGAATGAACGAGAGCTGGTTGCACATTAGGCTGTGA	1509
Qy	98	GlyValValLeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerProProArg	117	391	MetGlnCysMetGluThrHisLysLeuAspProAlaLysGluLeuProGlyTyrGlnIle	410
Db	493	GAAGCAGCTGTTTCTTGAAGCAAGATTCGTTGTCAG--AGGCTTCAAGAGAAAGA	549	411	LysGluGlnIleValSerLeuGluLysAspThrLeuGlnLeuAspLysGluMetGluGlu	430
Qy	118	AsnAsnValSer-----	121	1564	CAGAAACAGATTGGTGCAGCTAGAGAAG-----GCCAAAGCAGACAAGAGCGGAAACT	1617
Db	550	GATGCTGCTTCATTTGCCATTTGTAATGCTCGAGAAAGCAGAGAAATTTTCATCAGT	609	431	LysAlaArgSerLeuSerLeuMetGluGluAlaAlaLeuAlaLysArgMetTyrAsnGln	450
Qy	121	-----	121	1618	GAAGCAACA-----	1632
Db	610	GAATTCGCTACTTCTCTAATGCTGTAAGGAGGAGTCCAGGAGTGGAGGAGAAACCA	669	451	GlnIleLysArgProArgLeuSerProMetGluMetProValThrSerSerSerTyr	470
Qy	122	ValGluThrValThrValSerGlnProSerGlnGluIleVal-----ProGluThr	139	1633	CAACCCAAAGAGACCTCGTGCAAATGGTGTGGGATACGGTCCACGTGCACATAAC-----	1686
Db	670	GUGATACTTGTGCTCACTCGACTGAAGTAATGTGAAGATGTAAAACTTCTCTGATAAT	729	471	SerProIleTyrArgAspArgSerPheProSerGlnArgAspAspGlnAspGluIle	490
Qy	140	SerAsn-----LysProGluGlyArgMetCysGluLeuMetCysSer	154	1687	-----ATTCTTCGGATAAA-----	1701
Db	730	GGAAATGTGGAGTTGGTATCTTATCCAGATGGTTGTTAAACTATGCAAGAGATGGATGCT	789	491	SerAlaLeuValSerSerTyrLeuGlyProSerThrSerPheProHisArgSerArgArg	510
Qy	155	LysGlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGlu	174	1702	-----ACTTGCTATGCTAGAGTTGCTCACAGGTATCCACAATACGTTGATGAC	1749
Db	790	GCTGGACTTTCACAAATTCATATCTGATAACCGTAAGAACCTTGCTGTGAAGGATGAA	849	511	SerProGluTyrMetValProLeuProHisGlyLeuGlyArgSerValTyrAlaTyr	530
Qy	175	IleProSerAlaLeuLysLeuAlaLysGluProAlaLysPheValLeuAspCysIleGly	194	1750	CGACCACTTACATGTACCTCGCACCC-----	1782
Db	850	ATACCAAAATGCTTTAAGAGCTGCTCTAATGCTGCTGTAGTTTGTAGATTCTCTGAA	909	531	GluHisLeuAlaPro-----AsnSerTyrSerProGlyHisGlyHis	544
Qy	195	LysPheTyrLeuGln-----GlyArgAlaPheThrLysGluSerProMetSerSer	212	1783	ANTCATTGGCCCTCTCATGCACCATCGCAACATATAACATCTCTCCAGCCATGGCAAC	1842
Db	910	GGATTTTACTGTACAGAGTGTCAAATCAGGACGTGAAGAGGATGCTTAACCTATTGGGT	969	545	ArgLeuHisArg-----GlnTyrSerProSerLeuValHis	556
Qy	213	AlaArgGlnValSerLeuLeuLeuLeuSerPheLeu-----LeuMetProAspArgGly	231	1843	TACTTTGGAAATGGGTATCATGTACCAAGCCTCATATCTCCAC	1884
Db	970	CTTGGCCGAACCTGTATCATGCTGTGATGGAATGCTCTGTGTATTCCTCAGCAGCTCAGGT	1029	RESULT 5		
Qy	232	LysGlyLysValLysIleGluSerTrpIleLysAspGluAlaGluThrAlaAlaValAla	251	US-10-425-114-35432		
Db	1030	TTTGTGTTCTTAATGTGATTTTCAAGAGATATCAGGACAGCGCAAAAGCAGTTGCTGAAGAA	1089	; Sequence 35432, Application US/10425114		
Qy	252	TrpArgLysArgLeu-----MetThrGluGlyGlyLeuAlaAlaAlaGluLysMetAspAla	270	; Publication NO. US20040034888A1		
Db	1090	TGGAACCCAGGCTGATGCTCTTGACATGGATGTAGCAATGGGAATTCCTCGGAGCT	1149	; GENERAL INFORMATION:		
Qy	271	ArgGlyLeuLeuLeuValAlaCysPheGlyValProSerAsnPheArgSerThrAsp	290	; APPLICANT: Zhou, Yihua		
Db	1150	CATGCATTTTACAACTCTTCTAGTTTGAATTTGCTCTGTTTTGATGAGAGGAG	1209	; APPLICANT: Kovalic, David K.		
Qy	291	LeuLeuAspLeuIleArgMetSerGlySerAsnGluIleAlaGlyAlaLeuLysArgSer	310	; APPLICANT: Screen, Steven E		
Db	1210	TTATCTAGATTCGATTCGAATGCTGCTCGGCGCCCAAACTGCTGATTTTATGTCGTTTT	1269	; APPLICANT: Tabaska, Jack E		
				; APPLICANT: Cao, Yongwei		
				; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With		
				; FILE REFERENCE: 38-21(53313)B		
				; CURRENT APPLICATION NUMBER: US/10/425,114		
				; NUMBER OF SEQ ID NOS: 2003-04-28		
				; SEQ ID NO 35432		
				; LENGTH: 1931		
				; TYPE: DNA		

852	ATGGAGTCACTT	-----GGACAACTGCACCTTAATGATATA	887
236	-----LysIleGluSerTrpIleLysAspGluAlaGluThr	247	
888	ACTGTGTTCTTCTTTGGAGAGCACATGCTTACAACTAATATCATAGACGACAAAGACG	947	
248	AlaAlaValAlaIleArgLysArgLeu	261	
948	ATTGCATTTAAAGTGAAGTCCAAAGTTAGATAATCTTTGACATGTAGCATGGGAAC	1007	
262	LeuAlaAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPheGly	281	
1008	TGP-----CTTGAAGCTCATGCGTTCTTCAACTCTCTGGCAACCTTTGGT	1052	
282	ValProSerAsnPheArgSerThrAspLeuLeuAspLeuIle	297	
1053	ATTTCCTGCTGAATTCATCAATGAAGATGACTTTGTGCAAAATTCCTCCATATGTCAGTCGTCGT	1112	
298	SerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetVal	317	
1113	CGTCACAGCACTGAATCTTGTGATGCTGGGTGTCGAG-----AAATG	1160	
318	SerGlyIleValGluSerSerIleLysArgGlyMetHisIleGluAlaLeuGluMetVal	337	
1161	CCAGGTGTCAATTGAAGTCTCTGGTGAAGATGGAAGACCTATTGATGCAATTAATTTGGCT	1220	
338	TyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLys	357	
1221	TATGTGTTTGAGCTTACTGAACAAATTTGAACCAAGTACACCTCTCTTAAGCATATCTAAGG	1280	
358	MetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeuAlaPheLys	377	
1281	GATGTTTAAGAAA---ATGTCAACAGCGCAGGAATGTCAAAACTTCTCTGTGAGCACAGAAT	1337	
378	GluAlaAlaThrLysGlnLeuAlaValLeuSerSerValMetClnCysMetGluThrHis	397	
1338	GAGATGAATCAACGTGAGTTATCTGCCCTGAATCTGTGATTAAGTGCATTAAGAGCAC	1397	
398	LysLeuAspProAlaLysGluLeuProGlyTrpGlnIleLysGluGlnIleValSerLeu	417	
1398	AAACTTGAG-----CAGCAGTACCCT-----GTGGATCCACTT	1430	
418	GluLysAspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeu	437	
1431	CAGAAAGGGTTTTGCAGCTGAGAAACCAAGGACAGACAGAGGATGCGCTGT	1484	
438	MetGluGluAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLeu	457	
1485	-----GAAGCAGCA-----AAAGCCGAGTCCAAAGAGCGCTCGTCC	1520	
458	SerProMetGluMetProValThrSerSerSerTyrSerProIleTyrArgAspArg	477	
1521	ATTTGATCGGCCTT-----GCMAACCGTGTACAGGCTTTGCTGCACAAG	1565	
478	SerPheProSerGlnArgAspAspGlnAspGluIleSerAlaLeuValSerSerTyr	497	
1566	AGCTTCTACTCA-----	1577	
498	LeuGlyProSerThrSerPheProHisArgSerArgSerProGluTyrMetValPro	517	
1578	-----GCAACCCAGAGAGCATCCATTCAATCTTTCGAGAGAG-----	1616	
518	LeuProHisGlyGlyLeuGlyArgSerValTyrAlaTyrGlu---HisLeuAlaPro---	535	
1617	-----CAGTTTGTGTATGGCGCTGAGGTTTCATCTTCCTCCCATG	1655	
536	-----AsnSerTyrSerProGlyHisGlyHisArgLeu	546	
1656	ATGACCTCAGCTCTTACAGGATGCAACCTGCCATGGACCTTATTACGGTAATGGCTAC	1715	
547	HisArgGlnTyrSerProSerLeuValHisGlyGlnArgHisPro	561	
1716	CCAGTTTCAGTACACAGGCTCGTATATCTTCACTAATGAGACAGCA	1760	

[illegible]

Db 1790 GACTCTGGTGCATTTTCCTTACAGTGAACCTTTTCAGTGGCCCAAAACCAATTCACGTAC 1849
Qy 545 GLeuHisArgGlnTyrSer-----ProSerLeuValHisG1 557
Db 1850 ACTCAGAGGTCAGTGGCAGCATCATCAATTCAGCCCTTAAAGTCGTATTAGTGGC 1909
Qy 557 yGlnArgHis-----ProLeuGlnTyrSerProProleHisGlyGlnGlnG1 573
Db 1910 CCAGAGCAGCAGTGCAGAGCACCTATCCAGGTTATG---CCAGGCATCTGGCCCGCT 1966
Qy 573 nLeuProTyrGlyGlnArgVal-----TyrArgHisSerProSerGluGluAr 590
Db 1967 GCTTCACAGCAGTATGCAAAATATATATGATCGATATATCGCTCCCCACACCATAG 2026
Qy 590 gTyr 591
Db 2027 ATTC 2030
RESULT 7
US-10-425-114-5736
; Sequence 5736, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 5736
; LENGTH: 2300
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; FEATURE INFORMATION: Clone ID: 700471422_FLI
US-10-425-114-5736
Alignment Scores:
Pred. No.: 3,858-25 Length: 2300
Score: 325.00 Matches: 151
Percent Similarity: 41.80% Conservative: 109
Best Local Similarity: 24.28% Mismatches: 251
Query Match: 10.47% Indels: 111
DB: 12 Gaps: 25
US-09-890-475-1 (1-609) x US-10-425-114-5736 (1-2300)
Qy 19 LeuLeuGlnArgHisGlnSer---GluGlnArgArgArgGluLeuProLyHisLeuValGlu 37
Db 409 CTGTAGAGAGAGAGAGCATAGAGAAAGCTGAAGCTTGATGTTGATTGCA 468
Qy 38 ThrGluSerThrSerMetAspLeuThrGlyGlnSerLysGlnProGlnPheLeuLys 57
Db 469 GAGAAAGAGGCTAAAGTT-----TCTACAAAGAGCATGCTCCTCACTGAAT 513
Qy 58 SerIleAspGluLeu-----AlaAlaPheSerValAlaValGluThrPheLysArgGln 75
Db 514 GAGCTTCAGGAGCTAAGGATGCTGCTCTCTTCTAGCAGAGGTTGCGACAAAATAT 573
Qy 76 PheAspLeuGlnLysHisIleGlu-----SerIleGluAsnAlaIleAspSer 92
Db 574 AAGGTGGAGCTTGTCTGAGATCTTGAAGCAAGTGAAGCAAGCAAGCAAGCAAGTAACTACC 633
Qy 93 LysLeuGluSerAspGlyValValValValValValValValValValValValValVal 112
Db 634 TCAATCAATGACAACAATGATCTCTGCTTCAGAGGAGAAAT----- 675

Qy 113 LeuSerProProArgAsnAsnValSerValGluThrThrValThrValSerGlnProSer 132
Db 676 -----ACCCCTGCTAGTGGGTGGGTGAGGCA 702
Qy 133 GlnGluIleValPro---GluThrSerAsnLysProGluGlyGlyArgMetCysGluLeu 151
Db 703 TCTGAAGCTTTACTCTTGAGACTAAGCGGCCCTGTATTGAAGCAACTTTGTGAACAG 762
Qy 152 MetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeu 171
Db 763 ATGCACACCAAGGCTTCTGAAATTTCTTTTCAGAAAACTGTAGAAAACTTCTAGCTTT 822
Qy 172 MetGluGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLysPheValLeuAsp 191
Db 823 CGTGATGAATCTTCTGTTGCACTATAAATGTGCACTGATCTCCGCGCTTCGTACTTAAT 882
Qy 192 CysIleGlyLysPheTyrLeu-----GlnGlyArgArgAlaPheThrLysGluSerPro 209
Db 883 TCCTGGAGGCTTCTTCCACAGCAGCAAACTAATTCACCTGGGAGTAAACATAATGCC 942
Qy 210 MetSerSerAlaArgGlnValSerLeuLeuLeuGluSerPhe-----LeuLeu 226
Db 943 CTGAGGTCACGCGCAAGAGCTGCATCTTTTAATGAAGCTATAGCACCTGCCTAGGG 1002
Qy 227 MetProAspArgGlyLysGlyLysValLysIleGluSerTrpIleLysAspGluAlaGlu 246
Db 1003 ATGAAGGAGCTGGT---GGCGACGACCTTGGAGCTCTGAAATTAAGGAGCAAGCAAG 1059
Qy 247 ThrAlaAlaValAlaTrpArgLysArgLeuMetThrGluGlyGlyLeuAlaAlaGlu 266
Db 1060 GAAATTCGCGAAGAGTGAAGATTAAGTTA---GCTGAGATTGACCTGGATCTCTTAAT 1116
Qy 267 -----LysMetAspAlaArgGlyLeuLeuLeuValAlaCysPheGlyValProSer 284
Db 1117 GGTATTCTATGGAGGACAGGCTTCTCTGACGCTCTTACAACITTTTAATGTGATTCG 1176
Qy 285 AsnPheArgSerThrAspLeuLeuAspLeuLeuLeuLeuValAlaCysPheGlyValProSer 304
Db 1177 GTGCTTGCAGAGATGAACATATGCAAGATGCTAGTTGTCTCTCTGCTGCGACGAGCT 1236
Qy 305 GlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIleValGluSerSer 324
Db 1237 GCTGATCTCTGCTCTCTCTGCTTATGAGAAAGTACCAGGTATCATTCAGGAGTTG 1296
Qy 325 IleLysArgGlyMetHisIleGluAlaLeuGluMetValTyrThrPheGlyMetGluAsp 344
Db 1297 GTTAAGCGGCACAGGCAAAATGATGAGTTCATTTTCATCAAGCTTTTGGGCTTCGAG 1356
Qy 345 LysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGlu 364
Db 1357 ACATTCCTCCCTGCACCTCTGCTGAAGACATATGTTGAGAAACAAAGATACAAATTGA 1416
Qy 365 ArgAlaLysArgLysAlaGlnSerProLeuAlaPheLysGluAlaAlaThrLysGlnLeu 384
Db 1417 AAT---AATGGCGATGCAACTGCGATCTCATTAACGATGACCCCAAACTAGGAGCTA 1473
Qy 385 AlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeuAspProAlaLysGlu 404
Db 1474 ATTGCTTTGAGGCTGTGATTAAGTGTATGAGACATAAGCTT----- 1518
Qy 405 LeuProGlyTrpGlnIleLysGluGlnIleValSerLeuGluLysAspThrLeuGlnLeu 424
Db 1519 -----CAGAAGGAGTGTTCACCTGGACCTCTCCAGAAGCGTGTCTTCTGAGCTG 1566
Qy 425 AspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGluGluAlaAlaLeuAla 444
Db 1567 AAGCCCAAGGGTGAAGAAAGGCGCATCAAGT-----GATGCT 1602
Qy 445 LysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSerProMetGluMetProPro 464
Db 1603 GGGCGTACTTAT-----GCAAAGAGCTTCGAGGCGCTGGCATTTCAATTCCTAGG 1653
Qy 465 ValThrSerSerSerTyrSerProIleTyrArgAspArgSerPheProSerGlnArgAsp 484


```
Db 1654 AGACTGCCGGTCTGTTGGTTCAGCTCGCAGGCTCCATTTCCG----- 1701
Qy 485 AspAspGlnAspGluIleSerAlaLeuValSerTyrLeuGlyProSerThrSerPhe 504
Db 1702 -----GGCTATACCTGGCAGCGTGCTCCTGCG 1728
Qy 505 ProHisArgSerArg-----ArgSerProGluTyrMetValPro 517
Db 1729 CCGATGCCCTCCCGTGGTCTGCTCCAGTGCCTCCCGTGCCCT---CTTCTGCCCT 1785
Qy 518 LeuPro-----HisGlyGlyLeuGlyArgSerValTyr-----AlaTyr 530
Db 1786 CTTCCTGTAGATATAGGAGCTCTGACCGGTACCATTTACACACCAGCCCGCATAC 1845
Qy 531 GluHisLeuAlaProAsnSerTyrSer-----Pro-GlyHisGlyHisAr 545
Db 1846 GACTCTGGTGCATTTTCTGTCTTACAGTGAACCTTTCACTGCGCCCAAAACCATTCAGTAC 1905
Qy 545 GluHisArgGlnTyrSer-----ProSerLeuValHisG1 557
Db 1906 ACTCAGGGTCACTGGCAGCATCATCAATTCAAGCCCTATAAGTCGTTATGGTGGC 1965
Qy 557 yGlnArgHis-----ProLeuGlnTyrSerProProHisGlyGlnGlnG1 573
Db 1966 CCAGGAGCAGCAGTCAAGAGCACCTATCCAGGTATG---CCAGCGCATCTGGCGCGCT 2022
Qy 573 nLeuProTyrGlyIleGlnArgVal-----TyrArgHisSerProSerGluGluAr 590
Db 2023 GCTTCCAGCAGTATGCAAAATTATATGGATCCGCTATATCGCTCCCGCAACCATAG 2082
Qy 590 gTyr 591
Db 2083 ATTC 2086
```

RESULT 8

```
US-10-425-114-32984
; Sequence 32984, Application US/10425114
; Publication NO. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 32984
; LENGTH: 1780
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURES:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17048G12_FLI
US-10-425-114-32984
```

```
Alignment Scores:
Pred. No.: 1,32e-23 Length: 1780
Score: 309.50 Matches: 138
Percent Similarity: 41.27% Conservative: 89
Best local Similarity: 25.09% Mismatches: 223
Query Match: 9.97% Indels: 100
DB: 12 Gaps: 21
```

US-09-890-475-1 (1-609) x US-10-425-114-32984 (1-1780)

```
Qy 85 SerIleGluAsnAlaIleAspSerLysLeuGluSerAsnGlyValValLeuAlaAarg 104
Db 5 AGCAAGACAAAGGTAAGTACCTCAATCAATGACAAACATGCAATCTCGTGTTCAGAG 64
```

```
Qy 105 AsnAsnAsnPheHisGlnProMetLeuSerProProArgAsnAsnValSerValGluThr 124
Db 65 GAGAAAT-----ACC 73
Qy 125 ThrValThrValSerGlnProSerGlnGluIleValPro---GluThrSerAsnLysPro 143
Db 74 CTGCTAGTGGTTCGGCGAGGATCTGAAGCTTTACCTTGTGAGACTAAGCGCGCCCT 133
Qy 144 GluGlyGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyrIleTyrAla 163
Db 134 GTATTGAAGCAACTATGTGACAGATGACACCAAGAGGCTTCTGAAATTTCTTTTCAGAA 193
Qy 164 AsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAlaLeuLysLeuAlaLys 183
Db 194 AACTGTAGAAACCTTCTGATGATCACTTTCTGTGCACTAAATGTGCAACT 253
Qy 184 GluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeu-----GlnGlyArg 201
Db 254 GATCCTGCACGCTTCGTCTTAATTCCTGGAGGGTTCTTCCACACAGACCAACTAAT 313
Qy 202 ArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuLeu 221
Db 314 TCACCTGGAGTAAACATAATGCCCTTGAGGTCCAGCGCAAGAGCTGCATTTCTTTAATG 373
Qy 222 GluSerPhe-----LeuLeuMetProAspArgGlyLysGlyLysValLysIleGlu 238
Db 374 GAAGCTATAGCCTGCACCTAGGATGAAGAGCCTGGT---GGCGACGACCTTGGAGC 430
Qy 239 SerTyrIleLysAspGluAlaGluThrAlaAlaValAlaTyrArgLysArgLeuMetThr 258
Db 431 TCTGAAATTAAGGAGCAAGCAAGAAATTTGCCAAGAGAGTGAAGAGTAAGTTA---GCT 487
Qy 259 GluGlyGlyLeuAlaAlaGlu-----LysMetAspAlaArgGlyLeuLeuLeu 276
Db 488 GAGATTGACCTGATCTTCTAATGGCTATTCATGGAGGACAGAGCTTCTTCGACGCTT 547
Qy 277 ValAlaCysPheGlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuLeuArg 296
Db 548 CTTACAACCTTTAATGTGATTCGGTCTTGAAGAGATGAATATGCAAGATCGTAGTT 607
Qy 297 MetSerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMet 316
Db 608 GCTGTCTCTCGTCCAGAGAGAGCTGCTGATCTGCTCTCTCTTGGTCTTAATGAGAAA 667
Qy 317 ValSerGlyIleValGluSerSerIleLysArgGlyMetHisIleGluAlaLeuGluMet 336
Db 668 GTACCAAGTATCATTCAGGAGTGGTTAAGCGGCACAGCAAAATTCAGTCAGTTTCATT 727
Qy 337 ValTyrThrPheGlyMetGluAspLysPheSerAlaLeuValLeuThrSerPheLeu 356
Db 728 ATACAAAGCTTTTGGGCTTTCAGAGACATTCCTCCCTGACCTCTGCTGAAGACATATGT 787
Qy 357 LysMetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeuAlaPhe 376
Db 788 GAGGAACAAAGATACAAATTGAAAT---AATGGCATGCAACTCGCATCTCAATTAACG 844
Qy 377 LysGluAlaAlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThr 396
Db 845 GATGACCCCAAAATCTAGGAGCTAATTTGCTTTGAGGGCTGTGATTAAAGTGTATCGAAGAG 904
Qy 397 HisLysLeuAspProAlaLysGluLeuProGlyTyrGlnIleLysGluGlnIleValSer 416
Db 905 CATAAGCTT-----CAGAAGAGGTTCACCTTGGACCT 937
Qy 417 LeuGluLysAspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSer 436
Db 938 CTCAGAGCGGTGTTCTTCAGCTGAAGCCCAAGGGTGAAGAAAGGTCATCAAGT----- 991
Qy 437 LeuMetGluGluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgProArg 456
Db 992 -----GATGCTGGCGCTACTTAT-----GCAAGAGAGCGCTCGA 1024
```


QY 457 LeuSerProMetGluMetProProValThrSerSerSerTyrSerProIleTyrArgAsp 476
Db 1025 GSCCTGCAATTCATTTCTAGGAGACCTGCGGTTCTGTGTTCAGCTGCTCGCAGG 1084
QY 477 ArgSerPheProSerGlnArgAspAspGlnAspGluIleSerAlaLeuValSerSer 496
Db 1085 COTCATTTCCG-----GGC 1099
QY 497 TyrLeuGlyProSerThrSerPheProHisArgSerArg----- 509
Db 1100 TATACCTGGCAGCGTCTCTCGCGATGCGCTCCGTCGTCCTCGTCCGCTCC 1159
QY 510 ArgSerProGluTyrMetValProLeuPro-----HisGlyGlyLeuGlyArgSerVal 527
Db 1160 COTGCCCTCTCTCTGATAGATATGAGCTGCTGACCGGTACCAT 1216
QY 528 Tyr-----AlaTyrGluHisLeuAlaProAsnSerTyrSer----- 539
Db 1217 TACACACACAGCCCGCAGCATACGACTCGTGGTGCATTTTCTGTACGTGAACCTTC 1276
QY 540 -----Pro-GlyHisGlyHisArgLeuHisArgGlnTyrSer----- 551
Db 1277 AGCGCCCAAAACCATTCAGTACACTCCAGGTCAGTGGCAGCATCATACATTCAAGC 1336
QY 552 -----ProSerLeuValHisGlyGlnArgHis-----ProLeuGlnTyrSe 565
Db 1337 CCATATAAGTCGTTATGTTGGCCAGGAGCAGCAGTCAAGACCATTCAGTTAT 1396
QY 565 rProProlleHisGlyGlnGlnGlnLeuProTyrGlyIleGlnArgVal-----Ty 582
Db 1397 G-----CCGCGCATCTGGCGCGGCTCTTCCAGCAGCTATGCAAAATTATGGGATCCGTA 1453
QY 582 rArgHisSerProSerGluGluArgTyr 591
Db 1454 TATCGCCCTCCCAACCATAGATTTC 1481

RESULT 9

US-10-424-599-49586
; Sequence 49586, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5323)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 49586
; LENGTH: 2199
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: PAT_MRT3847_15784C.1
US-10-424-599-49586

Alignment Scores:

Prod. No.: 4,398-22 Length: 2199
Score: 297.00 Matches: 134
Percent Similarity: 38.65% Conservative: 89
Best Local Similarity: 23.22% Mismatches: 256
Query Match: 9.57% Indels: 98
DB: 12 Gaps: 16

US-09-890-475-1 (1-609) x US-10-424-599-49586 (1-2199)

QY 46 ThrIleGlyGlnSerLysGlnProGlnPheLeuLysSerIleAspGluLeuAlaPhe 65
Db 81 ACGATGGGGTCCATCCCGATCCAGCGAGTTGACGAGTTGACTCAG----- 128

QY 66 SerValAlaValGluThrPheLysArgGln----- 75
Db 129 ---CCGAGCTTCGAGAGTTCAGGCGCAAACTCCCTCATGACGAGTGCACCTCCTC 185
QY 76 PheAspAspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGlu 95
Db 186 TGAAGAGCTTCGAGCCACTTCTCTCCCTCCGAGCAGACCTCAACCAAAATCCGAA 245
QY 96 SerAsnGlyValValLeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerPro 115
Db 246 GCGCTCAAAACGCAAGATTGCGACCTCGACAACTACTTCGAGTCCCTCGCCTCCTC 305
QY 116 ProArgAsnAsnValSerValGluThrThrValThrVal----- 128
Db 306 GATCGCGCGAAACCTTCCTCGACGCGCTCCAGATCGCCCTCGGACGCTCGACAG 365
QY 129 -----SerGln 130
Db 366 CGCGCAGCGCTGCTCTCCGCTCTCCAGCGCGGACGACATCATCAACAGCTCC 425
QY 131 ProSerGlnGluIleValProGluThrSerAsnLysProGluGlyGlyArgMetCysGlu 150
Db 426 CCGACGCGTGGAGTGCAGACACCGCGCTCATCTCAAACTCAAGTCTTCTGCTC 485
QY 151 LeuMetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAlaLys 170
Db 486 CGCATGCGCGCATTCGGTCTTCGCTTCGTCAGCGCCCAAGAGAGCTCGACGC 545
QY 171 LeuMetGluGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLysPheValLeu 190
Db 546 CTACGCGCGAGATGCTGCTGCTGCGGAGTGTGCTATGATCCGCGCAAAATTCGTG 605
QY 191 AspCysIleGlyLysPheTyrLeuGlnGlyArgArgAlaPheThrLysGluSerProMet 210
Db 606 GAGCGCATCTCGAGGTGTTTCGCTGACAGAGAGGCGGCAAGCGCGCCAGCTG 665
QY 211 SerSerAlaArgGlnValSerLeuLeuIleLeuGluSerPheLeu-----LeuMet 227
Db 666 GCGTGGGCG-----TGCGCTCTTCTCTGGAGTGCCTGATTCCGGTCTGCTCGAC 716
QY 228 ProAspArgGlyLysGlyLysValIleGluSerTyrIleLysAspGluAlaGluThr 247
Db 717 CCGTCATCGGAAATTCGAGGCTGTGTGTGATCTCTATTGTGAAGGAGCAGCGCAG 776
QY 248 AlaAlaValAlaTyrPArgLysArgLeuMetThrGluGlyGlyLeuAlaAlaGluLys 267
Db 777 ATCGCAGAGCTTGGAGACCGCTCGAGGAGCGCGTGGGTGGAACCTGAAGACC 836
QY 268 MetAspAlaArgGlyLeuLeuValAlaCysPheGlyValProSerAsnPheArg 287
Db 837 CCGTGCCTCCACACTTCTTCGAGCAGCTGCTTACCTCGGATTGTCAAGAACGAGGAC 896
QY 288 SerThrAspLeuLeuAspLeuIleArgMetSerGlySerAsnGluIleAlaGlyAlaLeu 307
Db 897 TCGATTGTACCGGAGCTTGTATTGCTTCGCTTGGAGGAAACAGATGCCGAGCTC 956
QY 308 LysArgSerGlnPheLeuValProMetValSerGlyIleValGluSerSerIleLysArg 327
Db 957 GCGCTTTCGCTGCTCGCTCAGCAATGCTCATATGATTGAAGAGTTGATCAGCAA 1016
QY 328 GlyMetHisIleGluAlaLeuGluMetValTyrThrPheGlyMetGluAspLysPheSer 347
Db 1017 GGGCAGCAGCTTGATCGGTTCACTTTACATATGAGTGGGTCTGTGTGAAGATCCCT 1076
QY 348 AlaAlaLeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLys 367
Db 1077 CCGTTCCTCTGTTGAAGTCTTTTCTCAAG-----GATGCTAAG 1115
QY 368 ArgLysAlaGlnSerProLeu-----AlaPheLysGluAla 379
Db 1116 AAAGTTCGCGCTCTATTGGAAGATCTTAACATGACGCGCGAGCTCGTACCTAGCT 1175
QY 380 AlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeu 399

```
Db 1176 GCAAGGAAGGAGCTGCTCAGGCTGTGATTAAATGCAATGGAATCAAACTT 1235
Qy 400 AspProAlaLysGluLeuProGlyTrpGlnIleLysGluGlnIleValSerLeuGluLys 419
Db 1236 GAG-----GATGAGTTCACACAGAGAGATCTGAAGAGGAGCTTGACCACTGGAGAAG 1289
Qy 420 ---AspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuMet 438
Db 1290 GTGAAGACGGTGAAGACCGAAGAGG----- 1316
Qy 439 GluGluAlaAlaLeuAlaLysArgMetTyrAsnGlnIleLysArgProArgLeuSer 459
Db 1317 -----AAACCGGTGCTGCTCTCCCAATAGAGAACTCGAGCAGC 1358
Qy 459 ProMetGlu-----MetProValThrSerSerSerTyrSerProIleTyrArg 475
Db 1359 AACAGCAATGAGGTCCAAATGCCACCAAGCTGGGGGTTTGACTAATGCAATGTA 1418
Qy 476 AspArgSerPheProSerGln-----ArgAspAspAspGlnAspGluIleSer 491
Db 1419 TCA---TCTTTTCTCTGCTGCTCTACATTTGTCAGGCTCATCATCACAGGGCAATACCCA 1475
Qy 492 AlaLeuValSerSerTyrLeuGlyProSerThrSerPheProHisArgSerArgSer 511
Db 1476 GCTGCTCTTCCACCTTACCTTCCCTCCACCCACATGATGCGACAGAGTCCCCGACA 1535
Qy 512 ProGluTyrMetVal-----ProLeuProHisGlyLeuGlyArgSerValTyrAla 529
Db 1536 AATCCTTATGCTGCTTATTCACCTGAGCCGACCCGCTATTGCGGGTCTTACCCGGCA 1595
Qy 530 TyrGluHisLeuAlaProAsnSerTyrSerPro-----GlyHisGlyHisArgLeu 546
Db 1596 -----GCTCCCATGAATCTCTCTCTGATGATGGTGGCTATGGAATGTTTG 1643
Qy 547 HisArgGlnTyrSerProSerLeuValHisGlyGlnArgHisProLeuGln 563
Db 1644 GCTCCCACTTATCAGCAGGCTTACTACCGATAGAATGACAACCCCTTGAAG 1694
```

RESULT 10

```
US-10-424-599-95969
; Sequence 95969, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 95969
; LENGTH: 1812
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_57671C.1
US-10-424-599-95969
```

Alignment Scores:

Pred. No.:	6,95e-22	Length:	1812
Score:	294.00	Matches:	133
Percent Similarity:	38.69%	Conservative:	86
Best Local Similarity:	23.50%	Mismatches:	245
Query Match:	9.47%	Indels:	102
DB:	12	Gaps:	20

US-09-890-475-1 (1-609) x US-10-424-599-95969 (1-1812)

Qy 62 LeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAspLeuGlnLys 81

```
Db 28 CTGAACACTCTTCGACGCAAGAAAGAGATTCTCAAGAAAGCTACTACGATCTCCAAAGT 87
Qy 82 HisIleGluSerIleGluAsnAlaIleAsp-----SerLysLeuGluSerAsnGly 98
Db 88 CACTCTCTCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 147
Qy 99 ValValLeu---AlaAlaArgAsnAsnPheHis----- 109
Db 148 ACCTCCCTCCACACCTCTCTCTCACACGCTTCCACCTCTCTCTCTCTCTCTCTCTCTCTCTCT 207
Qy 110 -----GlnProMetLeuSerPro-----ProArgAsn 118
Db 208 CAACATATACCCACCATCTCTCTCCCTCCAAATACCTCAGCTTTCACCTCCCTCCACCCGAC 267
Qy 119 AsnValSerValGluThrThrValThrValSerGlnProSerGlnGluIleValProGlu 138
Db 268 CCATCATCGCAAAACGCGACCGCATTCGCCAAAATCCCTCGAACAATTCCTC----- 321
Qy 139 ThrSerAsnLysProGluGlyArgMetCysGluLeuMetCysSerLysGlyLeuArg 158
Db 322 -----ACGCTCTGCAACAACATGCGAAGAAAGGTTGAGG 357
Qy 159 LysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAla 178
Db 358 GACTAGCTCGGGACCAATTAAAGACAAGGCTGCAATCGAGGATACGCTTCGGAGTGCA 417
Qy 179 LeuLysLeuAlaLysGlu---ProAlaLysPheValLeuAspCysIleGlySerPheTyr 197
Db 418 CTCACAGAGTGCCTCTGATGCTGCCGATCGATCTTCTCGATTCGTCGTCGTCGTCGTCGTC 471
Qy 198 LeuGlnGlyArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSer 217
Db 472 GTTGTGGGCGCAATGTGGTGAAGATGACACGGAATTCGGCTCGAGGAAGAGACTTGC 531
Qy 218 LeuLeuIleLeuGluSerPheLeuLeuMetProAspArgGlyLysGlyLysValLysIle 237
Db 532 -----AGTTTTTTTGTCAAGCAGTTGAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 576
Qy 238 GluSerTrpIleLysAspGluAlaGluThrAlaAlaValAlaTrpArgLysArgLeuMet 257
Db 577 AGTTTAAAGGAGAAATTAAAGCCCAATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 636
Qy 258 ThrGluGlyGlyLeuAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuVal 277
Db 637 AGCGATGGT-----TGTGTGATGGTGTGGGGCCCATGCGCTTCTTGCATTTCGTG 687
Qy 278 AlaCysPheGlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMet 297
Db 688 GCGGCTTATGGCTCTCTCTGAATTGACCGTGCATGAGATTCTCACCCTCTCTCTTTATA 747
Qy 298 SerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetVal 317
Db 748 GCTGCTTCCAATGATGAGCTTCTCAGCTTACTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 807
Qy 318 SerGlyIleValGluSerSerIleLysArgGlyMetHisIleGluAlaLeuGluMetVal 337
Db 808 CCAGGCTTGTGCAAAAATTATTGACAGGAGCAAAACATATTCTGCGCTGTCAAGTTGT 867
Qy 338 TyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLys 357
Db 868 TTTGAGTTCAATCTTGCTCATAGATTCCACGATTCCTCCATTTTGGAGCTCATGTG--- 924
Qy 358 MetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeuAlaPheLys 377
Db 925 -----AATGAGTCTCAGAAACTTGTAAAGACTTCTGGAAGAGGGAAGTCACTTAGT 978
Qy 378 GluAlaAlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHis 397
Db 979 GAGATCACGCGAGAGAAATCCACTGCAATCGCGATTAAGGTTATTGAGATCAT 1038
Qy 398 LysLeuAspProAlaLysGluLeuProGlyTrpGlnIleLysGluGlnIleValSerLeu 417
```



```
QY 417 uGluLysAspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLe 437
Db 1298 AGAGAAG-----GTGAAGATGGTGAAGACCGAGAAACCAAGTGGCAGT 1345
QY 437 uMetGluGluAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLe 457
Db 1346 TCCTGCCAAT-----AAGAGAACTAGAGC 1369
QY 457 user-----ProMetGluMetProProValThrSerSerSerTyrSerPr 472
Db 1370 AAGCAACGGCAATGGAGTCCA-----ATGCCACCAAGCAAGCTGGGGTTGACTAA 1423
QY 472 oLleTyrArgAspArgSerPheProSerGln-----ArgAspAspAspGlnAs 488
Db 1424 TGCATATGTATCA---TCTTTCCTGCTGCTTCTTACATTTGTCCAGGTCCTCATCACACGG 1480
QY 488 pGluIleSerAlaLeuValSerSerTyrLeuGlyProSerThrSerPheProHisArgSe 508
Db 1481 GCATATCCAGCTCTCTTCCACCATACCTTCCCCACCCACACATGTACGGCAGCAGAAG 1540
QY 508 rArgArgSerProGluTyrMetVal-----ProLeuProHisGlyGlyLeuGlyArgSe 526
Db 1541 TCCCCAGCAAAATCTTATGCTTATTCACCCGAGCGGCACACAGCTATTGCAGGGTC 1600
QY 526 rValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSerProGlyHisGlyHisArgLe 546
Db 1601 TTACCCGGCAGTCCCTGAACTATCTCATGATATGGC---GGCTATGGAATGTTTT 1657
QY 546 uHisArgGlnTyrSerProSerLeuValHisGlyGlnArgHisProLeuGln 563
Db 1658 GGCTCCCACTTATCAGCAGGCTTACTACCGATAGAAATGCAACCTTTGAAG 1709
```

RESULT 12

```
US-10-425-114-34508
; Sequence 34508, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 34508
; LENGTH: 1772
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO1727806_FLI
US-10-425-114-34508
```

Alignment Scores:

Pred No.:	9.32e-19	Length:	1772
Score:	265.50	Matches:	114
Percent Similarity:	39.74%	Conservative:	68
Best Local Similarity:	24.89%	Mismatches:	206
Query Match:	8.55%	Indels:	71
DB:	12	Gaps:	12

US-09-890-475-1 (1-609) x US-10-425-114-34508 (1-1772)

```
QY 160 TyrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAlaLeu 179
Db 18 TTCGTGTGGCCCGCCGAGGAGGAGCGCGTGTGGGGCCGAGATGCCCGCGGCGTC 77
QY 180 LysLeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGln 199
```

```
Db 78 AAGCTCTCGCTGATCCCGCCCAAGTTCATGGACGCGCTGGCGAGCGTCTTCCCGCTC 137
QY 200 GlyArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeu 219
Db 138 GACCGCCGCGAGTCCGGAAACCCCGCCACCTGGCTGGCC-----TGCGTGTCTC 188
QY 220 IleLeuGluSerPheLeu-----LeuMetProAspArgGlyLysGlyLys 234
Db 189 ATCTCTGAGGCGCGCTGCTCCCGCTCCCGACACGGACCCCGACATCGGCCCGCGGC 248
QY 235 ValIleIleGluSerTrpIleLysAspGluAlaGluThrAlaAlaValAlaTyrArgLys 254
Db 249 CCGCTCGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCTG 308
QY 255 ArgLeuMetThrGluGlyGlyLeuAlaAlaGluLysMetAspAlaArgGlyLeuLeu 274
Db 309 GCCGCGGAGAAAGAGGGGGGCTCGAGGGCGCAAGCGCGCGCGCGCGCGCGCTTCTG 368
QY 275 LeuLeuValAlaCysPheGlyValProSerAsnPheArgSerThrAspLeuLeuAspLeu 294
Db 369 CAGCTCGTCCGCCACCTTTGCCGCTCCCGAGAGG---GAGGACAGGCCCTGTACCGCAGG 425
QY 295 IleArgMetSerGlySer---AsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeu 313
Db 426 ATTGCTGTCAGCTTCTCTGCGCGCGCGCGCGCGCGCGCTCGCGCTCACCTCGGACTC 485
QY 314 ValProMetValSerGlyIleValGluSerSerIleLysArgGlyMetHisIleGluAla 333
Db 486 GAGGAACAAATGGCCGATATCATTTGAGAACTGATTGTAAGAGGACAGCACTTGATGCT 545
QY 334 LeuGluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThr 353
Db 546 GTAATTTTCGCTTATCAGCTGGGCTTCAGGAGAGTTCCACCACTTCTCTTTTGAAG 605
QY 354 SerPheLeuLysMetSerLysGluSerPheGluArgAla-----LysArgLysAlaGln 371
Db 606 TCCTACCTGGAAGACTCTTAAGAAGACATCAACCGCTCTCAGATAATTAAGCACTAGC 665
QY 372 SerProLeuAlaPheLysGluAlaAlaThrLysGlnLeuAlaValLeuSerValMet 391
Db 666 AGCGGCCAATCAGGAGCAATGTGAACAAGAAAGAACAGCTGTGCACTCGAGCTGTGATA 725
QY 392 GlnCysMetGluThrHisLysLeuAspProAlaLysGluLeuProGlyTrpGlnIleLys 411
Db 726 AAGTGTGTGAGGATCGTAAACTAGAGCTGAGTTTCCACTGGAGGGT-----773
QY 412 GluGlnIleValSerLeuGluLysAspThrLeuGlnLeuAspLysGluMetGluGlyLys 431
Db 774 -----CTTCGGAAGCAACTTGAAGAACTAGAGAAAGCAAGCAAGCCGAGAG 818
QY 432 AlaArgSerLeuSerLeuMetGluGluAlaAlaLeuAlaLysArgMetTyrAsnGlnGln 451
Db 819 AAGAAGGACAGCATCA-----AGCGCTACAGTGGCGGCGAGCGGCGGCCCAACA 866
QY 452 IleLysArgProArgLeuSer-----ProMet-----460
Db 867 ACCAAGCGCATCCGTGCGAGCAGCTGGAGGCGCAATGCTCTCGCCAGGCGAGTGTCTC 926
QY 461 -----GluMetProValThrSerSerSerTyrSerProIle 473
Db 927 ACTAACAATGTTGGTGTCTTCTCTCCAGCTCCACCATGACCATGCTCTCCCTCC 986
QY 474 TyrArgAspArgSerPheProSerGlnArgAspAspGlnAspGluIleSerAlaLeu 493
Db 987 CACGGCTCATACGCCAACACCTCCCGTCCCGCACCTCATACGCAATACCTCCCATCC 1046
QY 494 ValSerSerTyrLeuGlyProSerThrSerPheProHisArgSerArgArgSerProGlu 513
Db 1047 CAGCGCTCATACGCCACAGCTCTCCCATCCCATGTCATCATACGACACCTCCCTCC 1103
QY 514 TyrMetValProLeuProHisGlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeu 533
```


Qy 475 ArgAspArgSerPheProSerGlnArgAspAspGlnAspGluIleSerAlaLeuVal 494
Db 1234 -----AACGCTTACGTC 1245
Qy 495 SerSerTyrLeuGluProSerThrSerPheProHisArgSerArgArgSerProGluTyr 514
Db 1246 TCCCTCTTCCACCGCTCCACACACATTCATCAGATCCCAAGCTCACTCACCTCAATAC 1305
Qy 515 MetValProLeuProHisGlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAla 534
Db 1306 GGTGTACCA-----GCATACACCACATCCCA 1332
Qy 535 ProHisSerTyr-----SerProGlyHisGlyHisArgLeuHisArgGlnTyrSer 551
Db 1333 CCTACCATCTACACACAGAGCGCTCCATAC-----CAATACATCA 1374
Qy 552 ProSerLeuValHisGly-----GlnArgHisProLeuGlnTyrSerProPheHis 569
Db 1375 CCCGAGCGGTTCTATGTTCTTACCAACCTCTCCGGTCAGTTAT---CCGACAGCATAT 1431
Qy 570 GlyGlnGlnGlnLeuProTyrGlyLe-----GlnArgValTyrArgHisSerProSer 587
Db 1432 GGTACTACTACGTCGGTGGTGGTCTCCACCCCTCCAGTTTACCATCTCCACCCGCAC 1491
Qy 588 GluGluArgTyrLeu 592
Db 1492 CACCACCACCATATT 1506

RESULT 14

US-09-938-842A-917
; Sequence 917, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SRIPL300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 917
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-917

Alignment Scores:

Pred. No.: 1,93e-15 Length: 1524
Score: 234.50 Matches: 117
Percent Similarity: 34.38% Conservative: 91
Best Local Similarity: 19.34% Mismatches: 222
Query Match: 7.55% Indels: 175
DB: 11 Gaps: 21

US-09-890-475-1 (1-609) x US-09-938-842A-917 (1-1524)

Qy 49 GlnSerLysGlnProGlnPheLeuLysSerIleAspGluLeuAlaAlaPheSerValAla 68
Db 34 AAGTCGTCTCAGCGAGTTCTTTGAG----- 60
Qy 69 ValGluThrPheLysArgGln-----PheAspAsp 78
Db 61 -----TTTCAGAGCAGCGCTCTCTAATGACGAGCTGCAATCTTCTATGAAAGAG 111

Qy 79 LeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsnGly 98
Db 112 CTCTCTGAGCATTTCCATCTCCATGAGCAAAACCTGATGAAGAAATCTGAAGCTCTGAGA 171
Qy 99 ValValLeuAlaAlaArgAsnAsnPheHisGlnProMetLeuSerProProArgAsn 118
Db 172 CAGATGATTGAACCCCTGATTAATCAACGCGAGAGCTCGATTGATTCTCAAGCATCGT 231
Qy 119 AsnValSerValGluThrThrValThrValSer----- 129
Db 232 GAGGTGACTATACACACAGCGCTCGAGATCGCGAGCGGAGAAAGTTGAGAACCGCTTAGA 291
Qy 130 -----GlnProSerGlnGluIleValProGluThrSerAsnLysPro 143
Db 292 GCGCGGTGGATTGCTCGTGAGAAAGCTAGAGACTGTGTGATGAAGATCCGAGGAGTC 351
Qy 144 GluGlyGly-----ArgMetCysGluLeuMetCysSerLys 155
Db 352 GACGACGGGACGGGCTTTTATCGGGCGTGAAGTCGCTTTGTTGAAGATGGACGCAAGA 411
Qy 156 GlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluLulle 175
Db 412 GGAATTTTGGGGTGTGTGATTTCGAGAGAGAGAAATTCGAGAAATCTCCGGTCACAGATT 471
Qy 176 ProSerAlaLeuLysLeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLys 195
Db 472 CCTGTGGCGTGTGATGTTGTGATCCGCCGAAGTTAGTACTTGAAGCTGTATCTGAG 531
Qy 196 PheTyrLeuGlnGlyArgAlaPheThrLysGluSerProMetSerSerAlaArgGln 215
Db 532 GTTTTTCGGTGGATAAAGAGGT---GGTGGAGAGAAAGTGAAGATGATTTTGGATGG 588
Qy 216 ValSerLeuLeuLeuGluSerPhe-----LeuLeuMetProAspArgGlyLys 232
Db 589 GCTGTGTGTGATCTGGAGAGTTGATTCGAGTTATGTTGATTCAGTGTATGGGAG 648
Qy 233 GlyLysValLysIleGluSerTrpIleLysAspGluAlaGluThrAlaAlaValaLtrp 252
Db 649 TCTCGGCTACTTGTAACTCCTAGCTTAAGGAGAAAGCTAAAGAGATGCTGAGACGTGG 708
Qy 253 ArgLysArgLeuMetThrGluGlyLeuAlaAlaGluLysMetAspAlaArgGly 272
Db 709 AAGGCTAGCTTGGAAAGAGAGAGAGGAGTGAAGATGTGAAACACCTGATTCATACG 768
Qy 273 LeuLeuLeuValAlaCysPheGlyValProSerAsnPheArgSerThrAspLeuLeu 292
Db 769 TTTCGCAACATCTTGTGACTTTTGGATTGGAAGAAG-----GATGATCTA 816
Qy 293 AspLeuIleArgMetSerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPhe 312
Db 817 GCTCTTTATAGG-----AAACTTGTGTGTGTTGAGTTCAGCTTGGCGTAAACAGATG 864
Qy 313 -----LeuValProMetValSerGlyValLeuGluSerSer 324
Db 865 CCAAGAGCTGCTGTTTACGTTGGTTGGCGACCAAAATCCCTGATATGATGAAGAGTTG 924
Qy 325 IleLysArgGlyMetHisIleGluAlaLeuGluMetValTyrThrPheGlyMetGluAsp 344
Db 925 ATAATCAGGGGACAAACAGCTTTCACGTTTCACTTTTGAAGTTGGTCTTGTACAC 984
Qy 345 LysPheSerAlaAlaLeuValThrSerPheLeuLysMetSerLysGluSerPheGlu 364
Db 985 TTGTTCCCTCCTGTTCTCTCAAGCTTATCTGAGGGAGTGAAGAAAGCTACAGCT 1044
Qy 365 -----ArgAlaLysArgLysAlaGlnSerProLeu 374
Db 1045 TTAATCACTGATGATTCACAAATCTCGCCGATCTCGGGGATCTGTTCTTCTCACTC 1104
Qy 375 AlaPheLysGluAlaAlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMet 394
Db 1105 GCTTAC-----CAATGTATG 1119


```

Qy 555 ValHisGly-----GlnArgHis-Pr 561
Db 935 ATCAG-GGAGCCATATGTTTACCCAGCTGAGGAGGTAGCCACCGTCAACGTCGGCATTC 993
Qy 561 oLeuGlnTyrSerProProlleHisGlyGlnGlnGlnLeuProTyrGly 577
Db 994 CATGGCCTACTCCACCCCGCCCATGAGCTACCCCTGCTGCCCCCTACGCG 1042

```

Search completed: February 28, 2004, 07:13:50
Job time : 558 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 28, 2004, 00:18:24 ; Search time 3455 Seconds
(without alignments)
5263.696 Million cell updates/sec

Title: US-09-890-475-1
Perfect score: 3104
Sequence: 1 MSNPFTVAQPTTANPLL.....RYLGLSNQSPRSNSLDPK 609

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-Q=/cgn2_1/USPFO.spool_p/US09890475/runat_24022004_135312_5304/app_query.fasta_1.775
-MCPDEL-frame+ p2n.model -DEV=xlp
-DB=EST -QFMT=fastap -SUFFIX=rest -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USPR=US09890475 @CGN 1 1 3549 @runat_24022004_135312_5304 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	742.5	23.9	704	28	BH491272	BH491272 BOHH16TR
C 2	696	22.4	706	28	BZ035667	BZ035667 oed83610.
C 3	687.5	22.1	781	28	BH479860	BH479860 BOGRJ75TF
C 4	663	21.4	552	9	AV541057	AV541057 AV541057
C 5	612	19.7	750	28	BH470959	BH470959 BOHLJ19TR
C 6	580	18.7	931	14	CK252208	CK252208 EST738845
C 7	556.5	17.9	943	14	CK252381	CK252381 EST736018
C 8	550.5	17.7	962	14	CK286517	CK286517 EST749239
C 9	543	17.5	635	28	B77833	B77833 T29E23TR TA
C 10	535.5	17.3	905	14	CK243005	CK243005 EST726642
C 11	528	17.0	913	14	CK294055	CK294055 EST756769
C 12	525.5	16.9	945	14	CK288931	CK288931 EST751653
C 13	523.5	16.9	684	14	CK836567	CK836567 BN45.049C
C 14	519.5	16.7	883	14	CK249078	CK249078 EST732715
C 15	518.5	16.7	904	14	CK243006	CK243006 EST726643
C 16	502.5	16.2	945	14	CK246861	CK246861 EST730498
C 17	494	15.9	840	14	CK286207	CK286207 EST748929
C 18	491.5	15.8	971	9	AJ568038	AJ568038 AJ568038
C 19	487	15.7	520	28	BZ512352	BZ512352 BOMSL50TF
C 20	486.5	15.7	785	13	BQ115269	BQ115269 EST760845
C 21	479	15.4	827	14	CK259125	CK259125 EST742762
C 22	467.5	15.1	557	10	BF050953	BF050953 EST436111
C 23	461	14.9	771	14	CK294226	CK294226 EST756940
C 24	456.5	14.7	903	14	CK252362	CK252362 EST735999
C 25	440	14.2	771	12	BI176101	BI176101 EST517184
C 26	435.5	14.0	688	28	BH984132	BH984132 oed42406
C 27	428	13.8	850	14	CK257335	CK257335 EST749572
C 28	423.5	13.6	671	12	BG890481	BG890481 EST516332
C 29	420.5	13.5	872	14	CK248078	CK248078 EST731715
C 30	420.5	12.6	896	14	CK258193	CK258193 EST741830
C 31	386.5	12.5	961	14	CK258053	CK258053 EST741690
C 32	385	12.4	685	10	AW738032	AW738032 EST739459
C 33	373	12.0	651	10	AW738612	AW738612 EST7340039
C 34	371	12.0	637	13	BQ118418	BQ118418 EST603994
C 35	353.5	11.4	540	10	BE471538	BE471538 EST416391
C 36	352.5	11.4	899	14	CK288916	CK288916 EST751638
C 37	352	11.3	500	13	BQ511650	BQ511650 EST7619065
C 38	321.5	10.4	2069	11	AY109393	AY109393 Zea mays
C 39	301	9.7	877	14	CK263597	CK263597 EST709675
C 40	295	9.5	544	10	BF053940	BF053940 EST439170
C 41	291	9.4	695	14	CK273406	CK273406 EST719484
C 42	287	9.2	861	29	CG962836	CG962836 MBH9N91TR
C 43	283.5	9.1	964	14	CK275053	CK275053 EST721131
C 44	279.5	9.0	430	12	BG097426	BG097426 EST462041
C 45	277.5	8.9	765	14	CB970934	CB970934 CAB10004_

ALIGNMENTS

RESULT 1
BH491272/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BH491272
BOHH16TR BOHB Brassica oleracea genomic clone BOHH16, genomic survey sequence.
BH491272
BH491272.1 GI:17699376
GSS.
Brassica oleracea
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 704)
 Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
 Whole genome shotgun sequencing of *Brassica oleracea*
 Unpublished (2001)
 Other GSSs: BOHBH167F
 Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.
 Location/Qualifiers
 1..704
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TOL000DH3"
 /db_xref="taxon:3712"
 /clone="BOHBH16"
 /note="Vector: pHOBI; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOBI using BstXI linkers"

ORIGIN
 Alignment Scores:
 Pred. No.: 5.98e-58 Length: 704
 Score: 742.50 Matches: 157
 Percent Similarity: 74.80% Conservative: 27
 Best Local Similarity: 63.82% Mismatches: 35
 Query Match: 23.92% Indels: 28
 Gaps: 7

US-09-890-475-1 (1-609) x BH491272 (1-704)
 QY 372 SerProLeuAlaPhe-----LysGluAlaAlaThrLysGlnLeuAlaVal 386
 Db 701 TCGTGTGGCTTTTATTTCTCAACAGAAATCGCGAACGAAAGACAGTGGATGCG 642
 QY 387 LeuSerSerValMetGlnCysMetGlnThrHisLysLeuAspProAlaLysGluLeuPro 406
 Db 641 TTATCATCAGTGAAGTGGTGTGAAGCTCACAGTTAGACCCAGTGAAGAAGTACCA 582
 QY 407 GlyTrpGlnIleLysGluGlnIleValSerLeuGlnLysAspThrLeuGlnLeuAspLys 426
 Db 581 GGGTGGCAGATCCAGAGCAATGCGAAGCTTGAAGAAGAGATTGTTTCAGCTCGACAA 522
 QY 427 GluMetGlu-----GluLysAlaArgSerLeuSerLeuMet 438
 Db 521 CAGATGGGAAGACGAGATCCATCAGTCGAATGGAGGAGCGCATCCATCAGTCTAAGG 462
 QY 439 GluGluAlaAlaLeuAlaLysArgMetTyrAsnGlnIleLysArgProArgLeuSer 458
 Db 461 GAGGAGCGGCAATTAGCGAGAGATTGTATTAACCAACAGATGAAAGCTCCAAAGTTGTCA 402
 QY 459 ProMetGluMetProProValThrSerSerSerTyrSerProIleTyrArgAspArgSer 478
 Db 401 GAAGGGAATGCCACCAACAGCTTCTTATCTCTTATCTCTATGATGATCCGCGACCAAGC 342
 QY 479 PheProSerGlnArgAspAspGlnAspGluIleSerAlaLeuValSerSerTyrLeu 498
 Db 341 TTCCCTAGTCACAGAGGAGATGCAGATGAATATCAGCTCTTGTGAGTAGTACCTC 282
 QY 499 GlyProSerThrSerPheProHisArgSer-----ArgArgSerProGluTyrMetVal 516
 Db 281 GGCCCATCAGCAGGTTTCTCTCATCGTCAAGTCTCAGGAGATCCCTCGAATATATGTT 222
 QY 517 ProLeuProHisGlyGluGlnValArgSerValTyrAlaTyrGluHisLeuAlaProAsn 536
 Db 221 -----CACCTGGTGGGTAGGAAGAGTGTCTCTGCTATGATCATCAGCCTCCAAT 168
 QY 537 SerTyrSerProGlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHis 556

167 TCTTATTCCTCG-----GTTTCAAGAGGTACTCTCCA-----GTTTCAAC 129
 QY 557 GlyGlnArgHisProLeuGlnTyrSerProIleHisGlyGlnGlnLeuProTyr 576
 Db 128 GGACAGAGACTTCTTCAAGAGTACTCTCTCCAGTTTCATGGCAACACCAATGCCATAT 69
 QY 577 GlyIleGlnArgValTyrArgHisSerProSerGluGlnArgTyrLeuGlyLeuSerAsn 596
 Db 68 GGTCTA-----TACAGACATTCCCATCTGT-GAAGATACTTGGCTTTGTCCAA 19
 QY 597 GlnArgSerProArgSer 602
 Db 18 CACAGGACTCTCTGTAAC 1

RESULT 2
 BZ035667/c
 LOCUS
 DEFINITION
 sequence.
 BZ035667
 BZ035667.1 GI:23612705
 GSS.
 Brassica oleracea
 Brassica oleracea
 Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 706)
 Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T.,
 Nash, W., Rabinowicz, P.D. and Wilson, R.K.
 Whole genome shotgun reads from *Brassica oleracea*
 Unpublished (2002)
 Contact: Richard K. Wilson
 Genome Sequencing Center
 Washington University School of Medicine
 Email: submissions@watson.wustl.edu
 Plate: oed83 row: e column: 10
 Seq primer: -28RPrOT reverse
 Class: shotgun
 High quality sequence start: 17
 High quality sequence stop: 551.
 Location/Qualifiers
 1..706
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /db_xref="taxon:3712"
 /clone_lib="B.Oleracea002"
 /note="Vector: pOTW13; Whole genome shotgun library from
 flowering buds. DNA was purified from a crude nuclear
 prep using *Brassica oleracea* TOL000DH3 buds provided by
 Thomas Osborn at the University of Wisconsin. Genomic
 DNA was provided by Pablo Rabinowicz (CSHL) and the
 shotgun library prepared at Washington University Genome
 Sequencing Center."

ORIGIN
 Alignment Scores:
 Pred. No.: 1.18e-53 Length: 706
 Score: 696.00 Matches: 146
 Percent Similarity: 75.11% Conservative: 26
 Best Local Similarity: 63.76% Mismatches: 35
 Query Match: 22.42% Indels: 23
 Gaps: 6

US-09-890-475-1 (1-609) x BZ035667 (1-706)
 QY 377 LysGluAlaAlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThr 396
 Db 657 AAACTGCCAAGCAAGCAAGCAGTGGATGCTTATCATCATGATGATGAAGTGT-TT-GAAGCT 599
 QY 397 HisLysLeuAspProAlaLysGluLeuProGlyTrpGlnIleLysGluGlnIleValSer 416

Db 598 CACAAGTTAGACCCAGTGAAGAAGTACACAGGCTGGCAGATCCCAAGAGCAATGGCGAAG 539
Qy 417 LeuGluLysAspThrLeuGlnLeuAspLysGluMetGlu----- 429
Db 538 CTTGAGAAAGAGATTGTTACGCTCGCAACACAGATGGAAGAGAGATCCATCAGTCGA 479
Qy 430 ---GluLysAlaArgSerLeuSerLeuMetGluGluAlaLeuAlaLysArgMetTyr 448
Db 478 ATGAGAGAGCGCATCCATCAGTCTAAGGAGAGAGCGGCAATTACCGAGAGATTGTAT 419
Qy 449 AsnGlnGlnIleLysArgProArgLeuSerProMetGluMetProProValThrSerSer 468
Db 418 AACCAACAGATGAACCTCCAGGTTCTCAGAAAGGAAATGCCACCAACAGCTTCTTAA 359
Qy 469 SerTyrSerProIleLysArgAspArgSerPheProSerGlnArgAspAspGlnAsp 488
Db 358 TCTTATTCTCTATGATCCGCGACCAAGCTTCCCTAGTCACAGAGAGGAGATGCAGAT 299
Qy 489 GluIleSerAlaLeuValSerSerTyrLeuGlyProSerThrSerPheProHisArgSer 508
Db 298 GAATATCAGCTCTTGTGAGTAGTACTCTGGCCCATCAGCAGGTTTCTCTATCGGTCA 239
Qy 509 -----ArgArgSerProGluTyrMetValProLeuProHisGlyLeuGlyArgSer 526
Db 238 AGTCTCAGGAGATCCCTGATATATGTT-----CCACCTGGTGGTTAGGAAGAAGT 185
Qy 527 ValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSerProGlyHisGlyHisArgLeu 546
Db 184 GTCTCTCGGTATGATCATCAGCTCCCAATCTTATTCTCCG-----GTT 140
Qy 547 HisArgGlnTyrSerProSerLeuValHisGlyGlnArgHisProLeuGlnTyrSerPro 566
Db 139 TCAAGAAGTACTCTCCA-----GTTCCGACAGAGACTTCTCTCAAGAGTACTCTCT 86
Qy 567 ProIleHisGlyGlnGlnGlnLeuProTyrGlyTleGlnArgValTyrArgHisSerPro 586
Db 85 CCAGTTCATGGCGCAACCAATCCATATGTTCTA-----TACAGACATTACCA 35
Qy 587 SerGluGluArgTyrLeuGlyLeuSer 595
Db 34 TCTGAATCCACCATGGAGCTCAGC 8

RESULT 3
BH479860
LOCUS
DEFINITION BOGRJ75TF BOGR Brassica oleracea genomic clone BOGRJ75, genomic survey sequence.
BH479860
BH479860.1 GI:17687964
GSS.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 781)
Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other GSSs: BOGRJ75TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source
1..781
/organism="Brassica oleracea"
/mol_type="genomic DNA"

/strain="T01000DH3"
/db_xref="taxon:3712"
/clone_lib="BOGRJ75"
/note="Vector: pHS01, Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHS01 using BstXI linkers"

ORIGIN

Alignment Scores:
Pred. No.: 8.52e-53 Length: 781
Score: 687.50 Matches: 144
Percent Similarity: 70.54% Conservative: 38
Best Local Similarity: 55.81% Mismatches: 57
Query Match: 22.15% Indels: 19
DB: 28 Gaps: 3
US-09-890-475-1 (1-609) x BH479860 (1-781)
Qy 28 ArgArgArgGluLeuProLysIleValGluThrGluSerThrSerMetAspIleThrIle 47
Db 19 CGACGAGAGAGAGCGCAGGCTACTGTGAAACCGTGCTACAAACATCGAAACACGATC 78
Qy 48 GlyGlnSerLysGlnProGlnPheLeuLysSerIleAspGluLeuAlaPheSerVal 67
Db 79 GAACAATCTAACGACCCCTCAGTTTGTGAATCCATCGCTGACTTAACCGCTTAGCACC 138
Qy 68 AlaValGluThrPheLysArgGlnPheAspAspLeuGlnLysHisIleGluSerIleGlu 87
Db 139 GCAGTGAACGCTTCAACGCGCTACGACGAACTGCAGAGCCACATGATTACATCAG 198
Qy 88 AsnAlaIleAspSerLysLeuGluSerAsnGlyValValLeuAlaAlaArgAsnAsnAsn 107
Db 199 AACGGGATCGACTCCCAATCTCAAACTAACGCGCATCGTCAAAATCGCC----- 246
Qy 108 PheHisGlnProMetLeuSerProArgAsnAsnValSerValGluThrThrValThr 127
Db 247 -----GCCGTGTGCTCTCCCGAAGAGTGGCTCTGGAGAAACAGCCACGGCG 294
Qy 128 Val-----SerGlnProSerGlnGluIleValProGluThrSerAsnLysProGlu 144
Db 295 ATTCGTGTCATCGCCGCCCAACAG-----AAGTCCGAA 330
Qy 145 GlyGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsn 164
Db 331 GCGGAGCGATTGTGCGAGTCGATGTGTAGCAAGAGCTCCGAGGTACATGTTCGTGAC 390
Qy 165 IleSerAspGlnAlaLysLeuMetGluGluProSerAlaLeuLysLeuAlaLysGlu 184
Db 391 ATATCTGAGAGAGCCAGCTAATCGAGAGCTTCTCGAGCGTTGAAGCTTGCCAAAGGAC 450
Qy 185 ProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGlyArgArgAlaPhe 204
Db 451 CCGGCGCAAGTTCGTGTGATTGCAATGGGAAGTTTACTTTCGAAAGGCGCAAGCGTTC 510
Qy 205 ThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuIleLeuGluSerPhe 224
Db 511 GCCACGACTCCCGCGCATCACCGAGAGAGTTTCTCTTCTTGTCTGAGTGTAT 570
Qy 225 LeuLeuMetProAspArgGlyLysValLysIleGluSerTrpIleLysAspGlu 244
Db 571 CTTCTGACGTTTGTCTCTGAAGAGAGAGAGTGTGTGTCTCTGTAAGATGAG 630
Qy 245 AlaGluThrAlaAlaValAlaTrpArgLysArgLeuMetThrGluGlyGlyLeuAlaAla 264
Db 631 GCGGAGCGCGTGTGTGTGCGTGGAGAGAGAGCTGGTGGGTGAAGATGTTGGGTGCA 690
Qy 265 AlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPheGlyVal 282
Db 691 GCGGAGCGTGTGATGCAAGGGGTTGCTTCTGTGTTGTTGTTGTTGTTGTTGTTGTT 744

RESULT 4

AV541057/c
LOCUS

AV541057 552 bp mRNA linear EST 07-SEP-2000

DEFINITION AV541057 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
 cDNA clone RZ159d09F 3', mRNA sequence.
 ACCESSION AV541057
 VERSION AV541057.1 GI:8702815
 KEYWORDS EST.
 SOURCE Arabidopsis thaliana (chale cress)
 ORGANISM Arabidopsis thaliana
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 552)
 Asanizu, E., Nakamura, Y., Sato, S. and Tabata, S.
 A large scale analysis of cDNA in Arabidopsis thaliana: Generation
 of 12,028 non-redundant expressed sequence tags from normalized and
 size-selected cDNA libraries
 JOURNAL DNA Res. 7 (3), 175-180 (2000)
 MEDLINE 20363093
 PUBMED 10907847
 COMMENT Contact: Erika Asanizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asanizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
 Location/Qualifiers
 1..552
 /organism="Arabidopsis thaliana"
 /mol_type="RNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="RZ159d09F"
 /tissue_type="roots"
 /clone_lib="Arabidopsis thaliana roots Columbia"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
 XhoI"

FEATURES
 source
 1..552
 /organism="Arabidopsis thaliana"
 /mol_type="RNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="RZ159d09F"
 /tissue_type="roots"
 /clone_lib="Arabidopsis thaliana roots Columbia"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
 XhoI"

ORIGIN

Alignment Scores:
 Pred. No.: 8,8e-51 Length: 552
 Score: 663.00 Matches: 122
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 21.36% Indels: 0
 DB: 9 Gaps: 0
 US-09-890-475-1 (1-609) x AV541057 (1-552)
 QY 488 AspGluIleSerAlaLeuValSerTyrLeuGlyProSerThrSerPheProHisArg 507
 DB 551 GATGAATATACCTCTTGATGATAGTACCTCGCCCGTCAACATCTTTCTCATCGC 492
 QY 508 SerArgArgSerProGluTyrMetValProLeuProHisGlyLeuGlyArgSerVal 527
 DB 491 TCAGAAGATCCCGGAATATATGTTCCACTTCCACATGTTGGGTAGGAAGAAGTGA 432
 QY 528 TyrAlaTyrGluHisLeuAlaProAsnSerTyrSerProGlyHisGlyHisArgLeuHis 547
 DB 431 TATGCATATGAACATCTGCGCCCAATTCATATCTCCAGGTCACGGACATAGACTTCAT 372
 QY 548 ArgGlnTyrSerProSerLeuValHisGlyGlnArgHisProLeuGlnTyrSerProPro 567
 DB 371 CGACAGTACTCTCCGCTTTGGTTTCACGGACAGACATCCACTACATGACTCTCTCTCCA 312
 QY 568 IleHisGlyGlnGlnGlnLeuProTyrGlyIleGlnArgValTyrArgHisSerProSer 587
 DB 311 APTCATGACACACACAGTACCATATGATATACAAAGGGTTACAGACATTCACCATCT 252
 QY 588 GluGluArgTyrLeuGlyLeuSerAsnGlnArgSerProArgSerProAsnSerSerLeuAsp 607
 DB 251 GAAGAAGATATTGGGTTTATCCAAATCAAGGTTCTCTCGCAGTAACTCATCTAGAC 192
 QY 608 ProLys 609
 DB 191 CCCAAA 186

RESULT 5
 BH470959/c
 LOCUS
 DEFINITION BOHLJ19TR BOHL Brassica oleracea genomic clone BOHLJ19, genomic
 survey sequence.
 ACCESSION BH470959
 VERSION BH470959.1 GI:17679070
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 REFERENCE 1 (bases 1 to 750)
 Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
 Whole genome shotgun sequencing of Brassica oleracea
 JOURNAL Unpublished (2001)
 COMMENT Other GSSs: BOHLJ19TF
 Contact: Chris Town

TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.

FEATURES
 Location/Qualifiers
 1..750
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO1000DH3"
 /db_xref="taxon:3712"
 /clone="BOHLJ19"
 /clone_lib="BOHLJ19"
 /note="Vector: pHS01; Site_1: BstXI; 2-3 kb sheared
 genomic DNA inserted into pHS01 using BstXI linkers"

ORIGIN

Alignment Scores:
 Pred. No.: 7,48e-46 Length: 750
 Score: 612.00 Matches: 135
 Percent Similarity: 70.22% Conservative: 23
 Best Local Similarity: 60.00% Mismatches: 34
 Query Match: 19.72% Indels: 35
 DB: 28 Gaps: 2
 US-09-890-475-1 (1-609) x BH470959 (1-750)
 QY 313 LeuValProMetValSerGlyIleValGluSerSerIleLysArgGlyMethHisIleGlu 332
 DB 666 ATTATCTATCGTCTAGGTATAGTTGATTTAAGTATCAAGCTGGAAGCATATTGAA 607
 QY 333 AlaLeuGluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeu 352
 DB 606 GCATTTGGGATGATTTATACCTT-GGATAGAGATAGTTTTCGGCTTCTCGCTCTTA 548
 QY 353 ThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSer 372
 DB 547 ACTTCATCTTAAGGATGAGCAAGGAGTCATTTAGAGGGCCAAACAGAAAGCTCAAGCA 488
 QY 373 ProLeuAlaPheLys----- 377
 DB 487 CCGATAGCATTTGT-ATGGCCCTTCTTAGCTTCATTCATGACTCTCTTTTTTTTTTCT 429
 QY 378 ----- 382
 DB 428 TTTCTGATCTTACCATTGTTGCCCTTTGTTCTCTTAACAGAAAGAGCCCAACCAAAAG 369
 QY 383 GlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeuAspProAla 402
 DB 368 TTTTACGTGCTTGTGTTATCATGATGAAGTGT-TTGGAGGCTCACAACCTTAGACCCAGAG 309

QY 403 LysGluLeuProGlyTrpGlnIleLysGluGlnIleValSerLeuGluLysAspThrLeu 422
 Db 308 AAAGAAGTACAGGATGCGAGATCAAGACAAATGATTAAAGTTGGAGAAAGACATTTCT 249
 QY 423 GlnLeuAspLysGluMetGluGluLysAlaArgSerLeuMetGluGluAlaAla 442
 Db 248 CAATCGACAAACAGATGCAAGGGAAGCAAGATCCATGATTATGAGGAAACAGCA 189
 QY 443 LeuAlaLysArgMetTyraGlnGlnIleLysArgProArgLeuSerProMetGluMet 462
 Db 188 TTGACGAAGAGATTGTATAACCAACAGATGAACAGTCCAAGTGTTCAGACATGGAATG 129
 QY 463 ProProValThrSerSerSerTyraProIleTyraArgAspArgSerPheProSerGln 482
 Db 128 CCACACAGACTTCCTCATCTTATCTTCTACCTACCTCCGCAAGCTTCCCTAGTAC 69
 QY 483 ArgAspAspGlnAspGluIleSerAlaLeuValSerSerTyraLeuGlyProSerThr 502
 Db 68 AGAGAC-----GATGAATATACAGCTTGTGAGTAGTTACCTCGAGCCTTCACCA 18
 QY 503 SerPheProHisArg 507
 Db 17 GGTTCCTCATCGG 3

RESULT 6
 CK255208 931 bp mRNA linear EST 12-DEC-2003
 LOCUS
 DEFINITION EST738845 potato callus cDNA library, normalized and full-length
 Solanum tuberosum cDNA clone POCU48 5' end, mRNA sequence.

ACCESSION CK255208
 VERSION CK255208.1 GI:39811864
 KEYWORDS EST.

SOURCE

Solanum tuberosum (potato)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE

1 (bases 1 to 931)
 Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.

AUTHORS

Generation of ESTs from potato callus tissue

TITLE

Unpublished (2003)

JOURNAL

Contact: Robin Buell

COMMENT

The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from TIGR via potato@tigr.org

Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

source

1..931

Location/Qualifiers

/organism="Solanum tuberosum"

/mol_type="mRNA"

/cultivar="Kennebec"

/db_xref="taxon:4113"

/clone="POCCJ48"

/tissue_type="callus"

/lab_host="DH10B-Tona"

/clone_lib="potato callus cDNA library, normalized and

full-length"

/note="vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;

supplier: RNA was isolated from Solanum tuberosum var.

Kennebec callus tissue grown on solid media."

ORIGIN

Alignment Scores:
 Pred. No.: 9,64e-43 Length: 931
 Score: 580.00 Matches: 125
 Percent Similarity: 59.56% Conservative: 65
 Best Local Similarity: 39.18% Mismatches: 106
 Query Match: 18.69% Indels: 24
 DB: 14 Gaps: 4

RESULT 7
 CK252381
 LOCUS
 DEFINITION EST736018 potato callus cDNA library, normalized and full-length
 Solanum tuberosum cDNA clone POCBU72 5' end, mRNA sequence.

ACCESSION CK252381
 VERSION CK252381.1 GI:39806317

US-09-890-475-1 (1-609) x CK255208 (1-931)
 QY 113 LeuSerProProArgAsnValSerValGluThrThrValThrValSerGlnProSer 132
 Db 5 CTCCACCGCCACCGCGCCGAG-ACGCTGATACACCAC-----CCTTCA 48
 QY 133 GlnGluIleValProGluThrSerAsnLysProGluGlyGlyArg----- 147
 Db 49 CCGGCGTCACACCGCTCAGCCTCGATCAACCTCCGAGAAAGAGACTCAATCT 108
 QY 148 -----MetCysGluLeuMetCysSerLysGlyLeuArgLysTyr 160
 Db 109 CCGGATTCCACAGCTCGAAACCTCTGCAAGACGATGCGCGCAAGGGACTTCGAAGTTAC 168
 QY 161 IleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAlaLeuLys 180
 Db 169 ATAGTATCGAACTCCAGAGAAACACACTCGCGGAAGAACTCACCAAGGCATTTGAA 228
 QY 181 LeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGly 200
 Db 229 CTTGCACCAATGCAGCAAGCTTGTACTTAGCTGTATGGGAGACTTCTTTGCGAAAGG 288
 QY 201 ArgArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeu 220
 Db 289 GGCAAAGCTTTTGTAAAGATGCACAGATGATGCCAAGGGAAGCATCTGCATTTGTT 348
 QY 221 LeuGluSerPheLeuLeuMetProAspArgGlyLysValLysIleGluSerTrp 240
 Db 349 TTGGAGTGTCTTCTTGTGATG-----GTTTCGATGTAATTTGATGAGGG 393
 QY 241 IleLysAspGluAlaGluThrAlaAlaValAlaTrpArgLysArgLeuMetThrGluGly 260
 Db 394 GTAAAGGAAGAGCTGCACAAGCAGCAGTAATATGGAGAAGAGTTTGTGTGTAAGA 453
 QY 261 GlyLeuAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPhe 280
 Db 454 GGTATAAGAAAGGCTAGTACATGATGCGCGGAGATTGCTATTCTTTATTTGGTGT 513
 QY 281 GlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySer 300
 Db 514 GGGATTCCACAATTTTTTACAAAGACGCTAGGAGCTTGAATTCGAGTCAGTAATATC 573
 QY 301 AsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyTyr 320
 Db 574 AGGGAGATTTCTACTGCATCAGGAGATCAAAATGTCTTAATGGAAAGATTCAGAAATA 633
 QY 321 ValGluSerSerIleLysArgGlyMetHisIleGluAlaLeuGluMetValTyrThrPhe 340
 Db 634 ATAGAGGCGCATGGTGAACATAAGATGCAACTGCTGTAGATGTTGCCCTATCTTT 693
 QY 341 GlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLys 360
 Db 694 GGAAGTTGAGAAAAATGTAGCCCTTTCGAAATTTTGTCACTATTTTGAAGAACTTAAA 753
 QY 361 GluSerPheGluArgAlaLysArgLysAlaGln--SerProLeuAlaPheLysGluAla 379
 Db 754 GAATCGCTGAGAAAAAAGAAATGCCATCGCATGGTTCACATGCTGTGTGAATGAAGCA 813
 QY 380 AlaThrLysGlnLeuAlaValLeuSerValMetGlnCysMetGluThrHisLysLeu 399
 Db 814 AACAGAGAGGAATTTGCTACTATAAATCTGCTATTGAATGTTTGAAGCCACCACATT 873
 QY 400 AspProAlaLysGluLeuProGlyTrpGlnIleLysGluGlnIleValSerLeuGlu 418
 Db 874 GATCCCTCAAGCTTATTTCACGATTCCGATTCAGTACAGAAATATAGCTTTGAG 930


```

QY 572 GlnGlnLeuProTyrGlyLeuGlnArgValTyrArgHisSerProSerGluGluArgTyr 591
Db 449 CAACAGTTACCATGTTATCAAAAGGGTTTACAGACATTCACCATCTGAAGAAAGATAT 390
QY 592 LeuGlyLeuSerAsnGlnArgSerProArgSerAsnSerSerLeuAspProLys 609
Db 389 TTGGGTTTATCAATCAAAAGGTTCTCTCGAGTAACATCATATTAGACCCCAAA 336

RESULT 10
CK243005 905 bp mRNA linear EST 12-DEC-2003
LOCUS Solanum tuberosum cDNA clone FOCA208 5' end, mRNA sequence.
DEFINITION CK243005
ACCESSION CK243005
VERSION CK243005.1 GI:39787192
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 905)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
Generation of ESTs from potato callus tissue
Unpublished (2003)
Other ESTs: EST726641 EST726643 EST726644
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
Clones can be requested from TIGR via potato@igr.org
Seq primer: ANT TAG GTG ACA CTA TAG.

FEATURES
source
location/Qualifiers
1..905
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="FOCA208"
/tissue_type="callus"
/lab_host="DH10B-10nA"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

ORIGIN
Alignment Scores:
Pred. No.: 1.18e-38 Length: 905
Score: 535.50 Matches: 118
Percent Similarity: 57.38% Conservative: 57
Best Local Similarity: 38.69% Mismatches: 107
Query Match: 17.25% Indels: 23
DB: 14 Gaps: 4

US-09-890-475-1 (1-609) x CK243005 (1-905)

QY 109 HisGlnProMetLeuSerProProArgAsnAsnValSerValGluThrVal 128
Db 1 CATCTCCGATGAGAAATCTGGCTGCTTCGCGTATCCAGACTCCCAACCCACCGG 60
QY 129 SerGln-----ProSerGlnGluLeuValProGluThrSerAsnLysPro 143
Db 61 CCGCAGACGGCTGATACCAACCCCTTCCCGCGGTCCACCACCGTTCGATCAAAACC 120
QY 144 GluGlyGlyArg-----MetCysGluLeu 151
Db 121 TCATCCGAGAAAGAAAGAACTCAAAATCTCGGATTCACAGCTCGAAACCTCTGCAAGC 180
QY 152 MetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeu 171

```

```

Db 181 ATGTCCGGCAGGAGCTTCGAAGTTACATAGTATCGAAATCCCGAGAAAAACACACTG 240
QY 172 MetGluGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLysPheValLeuAsp 191
Db 241 CGCGAAGAACTCACCAGGCAATTCGAATTCGACCAATTCGACCAAGCTTGACTTACG 300
QY 192 CysIleGlyLysPheThrLeuGlnGlyArgAsnGalaPheThrLysGluSerProMetSer 211
Db 301 TGTATGGGAGACTTCTTCGAAAGGGGCAAGCTTTTGATAGAAATGCACAGATGAT 360
QY 212 SerAlaArgGlnValSerLeuLeuIleLeuGluSerPheLeuLeuMetProAspArgGly 231
Db 361 GCCACAAGGGAAGCATCTGCATTTGTTGGAGTGTCTTCTTGTGATG----- 408
QY 232 LysGlyLysValLysIleGluSerTrpIleLysAspGluAlaGluThrAlaValAla 251
Db 409 ---GGTTTCGATGTAATTGATGAAGGGTAAAGGAGAGCTGCACAGCAGCAGTAATA 465
QY 252 TrpArgLysArgLeuMetThrGluGlyGlyLeuAlaAlaAlaLysMetAspAlaArg 271
Db 466 TGGAGAGAGAGGTTTGTGATGACGAGGTATAAGAAAGGCTAGTACATGATGCGCG 525
QY 272 GlyLeuLeuLeuValAlaCysPheGlyValProSerAsnPheArgSerThrAspLeu 291
Db 526 GGAATGCTATTGCTTATTGGTGTGTTGGGATTCACCAATTTTTCACCAATGAGGACGT 585
QY 292 LeuAspLeuIleArgMetSerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGln 311
Db 586 AGGAGCTTGATTCGAGTGAATATACAGGAGATTTCTACTGCATCAGGAGATCAAT 645
QY 312 PheLeuValProMetValSerGlyIleValGluSerSerIleLysArgGlyMetHisIle 331
Db 646 GTGCTAATGGAAGAGATTCAGAAATATAGAGGCGCATGCTGAAGCATAGGAGTT 705
QY 332 GluAlaLeuGluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuVal 351
Db 706 GATGCTGATAGTGTGCTTATCTTATCTTGGATTGAGAAATGAGCCCTTCGGAAT 765
QY 352 LeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLysAlaGln 371
Db 766 TTGTCATCATTTTATAGAAACTTAAAGAAATCGCTGAGAAAAAGAAATGCCAATCGCAT 825
QY 372 ---SerProLeuAlaPheLysGluAlaAlaThrLysGlnLeuAlaValLeuSerVal 390
Db 826 GGTTCATCATGCTGTTGTAATGAAGCAACAGAGGGAATTTGCTACTATGAATCTGTC 885
QY 391 MetGlnCysMetGlu 395
Db 886 ATTCGAATGTTGGAA 900

RESULT 11
CK294055 913 bp mRNA linear EST 15-DEC-2003
LOCUS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
DEFINITION EST756769 Nicotiana benthamiana mixed tissue cDNA library,
normalized, full-length Nicotiana benthamiana cDNA clone NBMC57 5'
end, mRNA sequence.
ACCESSION CK294055
VERSION CK294055.1 GI:39877075
KEYWORDS EST.
SOURCE Nicotiana benthamiana
ORGANISM Nicotiana benthamiana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 913)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,
Skaskawicz, B., Jin, H. and Baker, B.
Generation of EST sequences from Nicotiana benthamiana
Unpublished (2003)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org

```


Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

```

source
1..913
Location/Qualifiers
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone_lib="Nicotiana benthamiana mixed tissue cDNA
library, normalized, full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Nicotiana benthamiana
tissues that include callus, roots from liquid culture
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (Pseudomonas syringae pv tomato 12 hr;
Xanthomonas campestris pv campestris 12 hr, 18hr;
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
campestris pv vesicatoria 18hr). RNA was isolated from
these tissues and pooled in approximately equal molar
amounts."
ORIGIN
Alignment Scores:
Pred. No.: 5,928-38 Length: 913
Score: 528.00 Matches: 116
Percent Similarity: 60.50% Conservative: 54
Best Local Similarity: 41.28% Mismatches: 101
Query Match: 17.01% Indels: 11
DB: 14 Gaps: 3
US-09-890-475-1 (1-609) x CK294055 (1-913)
QY 126 ValThrValSerGlnProSerGlnGluLeuValProGluThrSerAsnLys-ProGluG1 145
DB 87 ATCACTACCGGAGCGCGGAGAGAAAGAAAGACTCAAGAAACCACTCCAAATCCACAGC 146
QY 145 yGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsn1 165
DB 147 CGAAACCCCTCTCGACAGCATGTGGCGAGGAGCTTCGAACTTATATATATATCCACT 206
QY 165 e-----SerAspGlnAlaLysLeuMetGluGluLeuProSerAlaLeuLysLeuAl 182
DB 207 GCCACTCCAGAGCAAGAAACACACACTGGCGAACCACTCCAGAGGTATTAAAGCTTGC 266
QY 182 alysgluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGlyArgAr 202
DB 267 ACCTAACGCGAGCAAGCTTGTACTCAGCTGCGTGGAGACTTTTACGCTAAAGGGGCAA 326
QY 202 gAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuLeuG1 222
DB 327 AGATCTTAATAAGGACACCGCAGATGATTCCTATAGGGAAGCATCCGCAATGGTTTGA 386
QY 222 uSerPheLeuLeuMetProAspArgGlyLysGlyLysValLysIleGluSerTrpIleIy 242
DB 387 GTGTCTTTTATTGATG-----GGTTGAACGAAATGTAGTAGTGGTTTC 431
QY 242 sAspGluAlaGluThrAlaAlaValAlaTrpArgLysArgLeuMetThrGluGlyGlyLe 262
DB 432 AAAGAGAGTGAACAGCGCGGTATATATGGAGAGAGATGGTTGTATGAAGAGGTAT 491
QY 262 uAlaAlaAlaGlyLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPheGlyVa 282
DB 492 AAAAAAGGCAATAAATAATGCATGCGCGGGTTTGTCTTATTGGGTGTGTGGGAT 551
QY 282 lProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySerAsnG1 302
DB 552 TCACAAAGCTTTTAGGAATGAGGATGTTAGGAGATTTGATTCGGAAGAGTAAACATCAGGG 611
QY 302 uileAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIleValG1 322

```

```

DB 612 GATTAAGAGTGTCTTAGGAGATCAAAATGTGTCATCGAAAGATTCAGAAATATAGA 671
QY 322 userSerIleLysArgGlyMetHieIleGluAlaLeuGluMetValTyrThrPheGlyMe 342
DB 672 GGGCATGGTGAAGCATAAGATGAAGTTGAGGCTGTGATGTTGCTTATCTATCTTTGGCGT 731
QY 342 tGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLysGluSe 362
DB 732 TGAGGAAAGATAGTCCCTTACAAATTTTGTCATCGTTTACATGATTTAAAGAATC 791
QY 362 rPheGluArgAlaLysArgLysAlaGlnSerProLeuAlaPheLysGluAlaAlaThrBy 382
DB 792 GCTGAAGAAAGAAATGGAATTCACATGCTTCACATGATGTTGT-GAATAAGCAACAA 850
QY 382 s---GlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeuAspPr 401
DB 851 GAGGGAATTTGCTACTATGAAATCTGTCAATGTTTGGAGGCCCATACGTTGATTC 910
QY 401 o 401
DB 911 C 911

```

RESULT 12

CK288931

LOCUS

DEFINITION

EST751653 Nicotiana benthamiana mixed tissue cDNA library,

normalized, full-length Nicotiana benthamiana cDNA clone NBMK18 5'

end, mRNA sequence.

CK288931 GI:39866931

EST.

KEYWORDS

SOURCE

ORGANISM

Nicotiana benthamiana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

asterids; lamids; Solanales; Solanaceae; Nicotiana.

1 (bases 1 to 945)

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,

Skaskawicz, B., Jin, H. and Baker, B.

Generation of EST sequences from Nicotiana benthamiana

Unpublished (2003)

Other ESTs: EST751654

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from TIGR via potato@tigr.org

Seq primer: ATT TAG GTG ACA CTA TAG.

Location/Qualifiers

1..945

/organism="Nicotiana benthamiana"

/mol_type="mRNA"

/db_xref="taxon:4100"

/clone="NBMK18"

/tissue_type="abiotic and biotic stress-treated leaves,

callus tissue and root tissue"

/lab_host="DH10B-Tona"

/clone_lib="Nicotiana benthamiana mixed tissue cDNA

library, normalized, full-length"

/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;

supplier: RNA was isolated from Nicotiana benthamiana

tissues that include callus, roots from liquid culture

grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),

cold-stressed leaves (5 C 3 hr, 6hr), and pathogen

challenged leaves (Pseudomonas syringae pv tomato 12 hr;

Xanthomonas campestris pv campestris 12 hr, 18hr;

Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas

campestris pv vesicatoria 18hr). RNA was isolated from

these tissues and pooled in approximately equal molar

amounts."

ORIGIN

Alignment Scores:		Length:	945
Pred. No.:	1.07e-37	Matches:	116
Score:	525.50	Conservative:	57
Percent Similarity:	61.13%	Mismatches:	100
Best Local Similarity:	40.9%	Indels:	11
Query Match:	16.9%	Gaps:	3
DB:	14		
US-09-890-475-1 (1-609) x CK288931 (1-945)			
QY	126	ValThrValSerGlnProSerGlnGluLeuValProGluThrSerAsnLys-ProGluG1	145
DB	108	ATCATACCGGCGCGCGGAGAAAGAACTCAAGAAACCACTTCCAAATCCACAGCT	167
QY	145	YGLYArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsnI1	165
DB	168	CGAAACCCCTCTCAAGACGATGTGGCAAGGACTTCGAATATATATATATCGCAACT	227
QY	165	e-----SerAspGlnAlaLysLeuMetGluGluLeuProSerAlaLeuLysLeuAl	182
DB	228	GCACCTCCAGAGCAAGAAACACACTGCGGCAACACTCCACAGAGTATTAAAGCTTGC	287
QY	182	ALysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGlyArgA	202
DB	288	ACCTAACGCGGCGGAGCTGTACTCAGCTGCGTGGGAGACTTTTACGCTAAAGGGGCAA	347
QY	202	gAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuLeuG1	222
DB	348	AGATCTTAATAGGACACGACGATGATTCCTATAGGAGAGCATCGCATGTGTTTGA	407
QY	222	uSerPheLeuLeuMetProAspArgGlyLysValLysIleGluSerTyrIleLys	242
DB	408	GTGTCTTTTATTGATG-----GGTTGAACGAAATGATGAAGTGGTTTC	452
QY	242	shpGluAlaGluThrAlaAlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla	262
DB	453	AAAGAAAGCTGTAACAGCGCGGTATATGGAGAGAGATTTGGTGTATGAAGAGATAT	512
QY	262	uAlaAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuLeuValAlaCysPheGlyVa	282
DB	513	AAAAAGGCAATAAATATGATGCGCGGGTTTGTCTGTTGCTTATTGGGTGTTTGGGAT	572
QY	282	lProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySerAsnG1	302
DB	573	TCCACAAGCTTTTGGAAATGAGATGTTAGGAGATTTGATTCGGAAGAGTAACATCAGGGG	632
QY	302	uIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIleValG1	322
DB	633	GATTAAGTGCTCTTAGGAGATCAATGTGCTCATCGAAAGATTCAGAAATAATAGA	692
QY	322	uSerSerIleLysArgGlyMetHisIleGluAlaLeuLysMetValTyrThrPheGlyMe	342
DB	693	GGGCATGTTGAAGCATAAGATGGAAGTGTGAGCTGTGATGTTGCTTATACCTTTGGCGT	752
QY	342	tGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLysGluSe	362
DB	753	TGAGGAAAAGTTAGCCCTTACAAATTTTGCTATGCTTTTACATGAATTTAAAGATTC	812
QY	362	rPheGluArgAlaLysArgLysAlaGln---SerProLeuAlaPheLysGluAlaAlaTh	381
DB	813	GCTGAAGAAAAGGAAATGGAATACATGATGTTTGCATGATGTTGTGAATAAAGCAACAA	872
QY	381	rLysGlnLeuAlaValLeuSerValMetGlnCysMetGluThrHisLysLeuAspPr	401
DB	873	GAGGGAATGCTACTATGAATCTGTCTCAATGTTTGGAGCC-CATAACGTTGATCC	931
QY	401	oAlaLys 403	
DB	932	CTCAAG 938	
RESULT 13			
CD836567			
LOCUS			
684 bp mRNA linear EST 10-JUL-2003			
DEFINITION			
BN45_049C12F020103 BN45 Brassica napus cDNA clone BN45049C12, mRNA			
sequence.			
ACCESSION			
CD836567			
VERSION			
CD836567.1			
GI:32518507			
KEYWORDS			
EST.			
SOURCE			
Brassica napus (rape)			
ORGANISM			
Brassica napus			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.			
REFERENCE			
1 (bases 1 to 684)			
Genoplatte, a major partnership french program in plant genomics			
Genoplatte.			
Unpublished (2003)			
Contact: Genoplatte			
Genoplatte			
92, rue Henri Rochefort 91025 EVRY CEDEX France			
Tel: 33 1 69 47 54 00			
Fax: 33 1 69 47 54 10			
This sequence has been generated in the framework of the french			
plant genomics programme 'Genoplatte' (http://www.genoplatte.com			
and http://genoplatte-info.infobiogen.fr).			
FEATURES			
Location/Qualifiers			
1..684			
Source			
/organism="Brassica napus"			
/mol_type="mRNA"			
/cultivar="Jet Neuf"			
/db_xref="taxon:3708"			
/clone="BN45049C12"			
/tissue type="seed"			
/clone_lib="BN45"			
ORIGIN			
Alignment Scores:			
Pred. No.:	9.58e-38	Length:	684
Score:	523.50	Matches:	110
Percent Similarity:	66.98%	Conservative:	32
Best Local Similarity:	51.89%	Mismatches:	49
Query Match:	16.87%	Indels:	21
DB:	14	Gaps:	3
US-09-890-475-1 (1-609) x CD836567 (1-684)			
QY	16	AlaAsnProLeuLeuGlnArgHisGlnSerGluGlnArgArgGluLeuProLysIle	35
DB	96	TCATCGCGGACCATACACCGG-----	116
QY	36	ValGluThrGluSerThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPhe	55
DB	117	---GGAAACCGGGCTACAAACACGGAATCAGCATCGAACAAATCTAACCATCTCAATTT	173
QY	56	LeuLysSerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGln	75
DB	174	TTGAATTCGATCGACGATTTAACTGCGTTTTTCAGTCGATGACGCTTCAACCGCAC	233
QY	76	PheAspAspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGlu	95
DB	234	TACGACGACTTGCAAGCCCTCATGATTACATCAAGAACGCCATTGACTCGACTCTCAAG	293
QY	96	SerAsnGlyValValLeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerPro	115
DB	294	AGCAAAAGCATACCGCGGAGTCTCCCTCTCCGATCGCAGTCT-----	338
QY	116	ProArgAsnAsnValSerValGluThrThrValThrValSerGlnProSerGlnGluIle	135
DB	339	CCACGAAACGATGCTTCGGGAGAAACGGTTGCTGCCACAAATCGCCGCAAGAGACT	398
QY	136	ValProGluThrSerAsnLysProGluGlyGlyArgMetCysGluLeuLeuMetCysSerLys	155
DB	399	TGTGAGACAGTAGCGGAGAGGTGGAG-----CGATTGTCGAGTGTGATGTGAGCAA	452
QY	156	GlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIle	175

Seq primer: CAG GAA ACA GCT ATG ACC.

FEATURES

Location/Qualifiers

1..904

/organism="Solanum tuberosum"

/mol_type="mRNA"

/cultivar="kennebec"

/db_xref="taxon:4113"

/clone="POCA208"

/tissue_type="callus"

/lab_host="DH10B-TonA"

/clone_lib="potato callus cDNA library, normalized and

full-length"

/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;

supplier: RNA was isolated from Solanum tuberosum var.

Kennebec callus tissue grown on solid media."

ORIGIN

Alignment Scores: 4,398-37 Length: 904

Score: 518.50 Matches: 118

Percent Similarity: 57.13% Conservative: 57

Best Local Similarity: 38.56% Mismatches: 107

Query Match: 16.70% Indels: 24

DB: 14 Gaps: 4

US-09-890-475-1 (1-609) X CK243006 (1-904)

QY 109 HisGlnProMetLeuSerProProArgAsnValSerValGluThrThrValThrVal 128
 DB 2 AATCTCCGAGAACTGGCTGGCTCTGCGCTATCCGAGCTCCACGCCACCCGCG 61
 QY 129 SerGln-----ProSerGlnGluIleValProGluThrSerAsnLysPro 143
 DB 62 CCGCAGACGGCTGATACACCCCTTCACCGGGCTCACACCGCTCGACTCGATCAACC 121
 QY 144 GluGlyGlyArg-----MetCysGluLeu 151
 DB 122 TCATCCGAAGAAAGAGACTCAAAATCTCCGGATTACAGCTCGAAACCTCTGCAAGACG 181
 QY 152 MetCysSerLysGlyLeuArgLysTrpIleTyAlaAsnIleSerAspGlnAlaLysLeu 171
 DB 182 ATGTGGCGGCAAGGAGCTTCGAAGTTACATAGTATCGAACCTCCAGAGAAACACACTG 241
 QY 172 MetGluGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLysPheValLeuAsp 191
 DB 242 CGCGAAGAACTCACCAGGCACTTGAACCTTGACCAATGACAAAGCTTGTACTTAGC 301
 QY 192 CysIleGlyLysPheTyLeuGlnGlyArgAlaPheThrLysGluSerProMetSer 211
 DB 302 TGTATGGGAGACTTCTTCGAAAGGGGCAAGCTTTTGATAGAATGACAGATGATT 361
 QY 212 SerAlaArgGlnValSerLeuLeuLeuGluSerPheLeuLeuMetProAspArgGly 231
 DB 362 GCCACAGGGAAGCATCTGCATGGTTGGTGGAGTGTCTTCTTGTATG----- 409
 QY 232 LysGlyLysValLysIleGluSerTrpIleLysAspGluAlaGluThrAlaAlaValAla 251
 DB 410 ---GGTTTCGATGTAATGATGAAGGGGTAAAGGAAGAGCTGCACAGCAGCAATA 466
 QY 252 TrpArgLysArgLeuMetThrGluGlyLeuAlaAlaAlaGluLysMetAspAlaArg 271
 DB 467 TGGAGAAAGAGGTTTGTATGATGACGAGGTATAGAAGGCTAGTACATGGATCGCGG 526
 QY 272 GlyLeuLeuLeuValAlaCysPheGlyValProSerAsnPheArgSerThrAspLeu 291
 DB 527 GGATTCGATTGCTTATGGGTGTTTGGGATTCACAAATTTTTCACAAATGAGGACGTT 586
 QY 292 LeuAspLeuIleArgMetSerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGln 311
 DB 587 AGGAGCTTGATTCGATGAGTATATACAGGAGATTTCTACTGCTCAGGAGATCAAT 646
 QY 312 PheLeuValProMetValSerGlyIleValGluSerSerIleLysArgGlyMethIle 331

DB 647 GTCTAATCGAAAGAGATTCCAGAAATAATAGAGGGCATGGTGAAGCATAGAATGAAGTT 706
 QY 332 GluAlaLeuGluMetValTyThrPheGlyMetGluAspLysPheSerAlaAlaLeuVal 351
 DB 707 GATGCTGTAGATGTTGCCTATACTTTTGAGTTGAGGAAATAATGTAGCCCTTGCAGAAAT 766
 QY 352 LeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLysAlaGln 371
 DB 767 TTGTCATCATTTTTTAGAGAACTTAAAGAAATCGCTGAGAGAAAGAAATGCCAATCGCAT 826
 QY 372 ---SerProLeuAlaPheLysGluAlaAlaThrLys-GlnLeuAlaValLeuSerSerVa 390
 DB 827 GGTTCATGCTGTTTGTGAATGAGCAACAAGAGGGGATTTGCTACTATGAATCTGT 886
 QY 390 lMetGlnCysMetGlu 395
 DB 887 CATTGAATGTTGGAA 902

Search completed: February 28, 2004, 05:48:12
 Job time : 3467 secs